ESSAYS IN PUBLIC AND REGULATORY ECONOMICS

Dissertation

for the Faculty of Economics, Business Administration and Information Technology of the University of Zurich

to achieve the title of

Doctor of Economics

presented by

Ilja Neustadt

from the Russian Federation

approved at the request of

Prof. Dr. Peter Zweifel

Prof. Dr. Reinhard Madlener

The Faculty of Economics, Business Administration and Information Technology of the University of Zurich hereby authorises the printing of this Doctoral Thesis, without thereby giving any opinion on the views contained therein.

Zurich, 27.10.2010

the Dean: Prof. Dr. Dr. Josef Falkinger

Preface

In writing this dissertation, I have enjoyed the help and support of several people whom I would like to thank.
First of all, I am grateful to Prof. Dr. Peter Zweifel, my thesis supervisor, for his guidance and ongoing support. In particular, I am grateful for his introducing me to the methodology of discrete choice experiments, which became important in dealing with questions analyzed in Chapters 3 to 5. Furthermore, I want to thank Prof. Dr. Reinhard Madlener for co-supervising this dissertation and his introducing me to the field of energy economics.
A substantial part of the research leading to Chapters 2 to 5 was conducted within the framework of the project "Redistribution in Switzerland: Evidence and Guidance for Policy", supported by the Swiss National Science Foundation (SNF).
In addition, I would like to thank my former and current work colleagues at the Socioeconomic Institute: Sule Akkoyunlu, Michael Breuer, Patrick Eugster, Boris Krey, Philipp Morf, Maurus Rischatsch, Adrian Rohner, Micha Ruflin, Felix Schläpfer, Johannes Schoder, Yves Schneider, Michèle Sennhauser, Georgios Sismanidis, Harry Telser, Maria Trottmann, and Philippe Widmer. Moreover, I want to thank the whole team of the library for their support.
Finally, many thanks go to my family and wife Rachael for their continuous support and encouragement during my PhD studies.

Ilja Neustadt, Zurich, October 2010

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This dissertation deals with two topics in the fields of public and regulatory economics: preferences for income redistribution (Chapters 2 to 5) and regulation of markets for renewable energy (Chapters 6 and 7). While dealing with seemingly different topical issues, the main focus of the dissertation is on the analysis of public policy with regard to allocative efficiency (Chapters 6 and 7) and redistributive justice (Chapters 2 to 5). The first three essays (Chapters 2 to 4) provide a positive analysis with their focus on experimental measurement of preferences for redistribution and empirical explanation of their determinants. By way of contrast, the next three essays deal with normative questions. In particular, Chapter 5 derives policy implications as to the sustainability of the Swiss welfare state based on results from the previous chapters. Further, Chapters 6 and 7 provide an analysis of optimal policy instruments for the regulation of markets for renewable energy.

Chapter 2, "Why does the amount of income redistribution differ between the United States and Europe? The Janus face of Switzerland", is a result of an extensive review of literature on demand for income redistribution. In this essay, economic, political, and behavioral determinants of the amount of public social expenditures elaborated in the literature are reviewed. By using empirical data from the majority of OECD countries, this chapter compares the levels of income redistribution in the United States, the European Union, and Switzerland and relates them to the possible determinants. Lying in between the two poles, Switzerland provides unique evidence about the relative merits of competing hypotheses. It tips the balance against the economic explanations, which are based on income inequality and social mobility and predict more rather than less income redistribution in the United States compared to the EU. It only weakly supports the political and institutional explanations linking proportional representation to redistribution. However, behavioral explanations receive a good deal of support from the case of Switzerland, a country that shares with the United States the belief that effort rather than luck and connections determine economic success. In this way, the Janus face of Switzerland helps to explain the difference in the amount of public social spending in the United States and the European Union.

The next three essays, presented in Chapters 3 to 5, are an outcome of a Discrete Choice Experiment (DCE) performed in 2008 in Switzerland. Through this DCE, preferences for income redistribution in Switzerland were elicited. In addition to the amount of redistribution as a share of GDP, attributes also included its uses (working poor, the unemployed, old-age pensioners, families with children, people with ill health) and nationality of beneficiary (Swiss, Western European, others).

In Chapter 3, "Economic well-being, social mobility, and preferences for income redistribution: evidence from a discrete choice experiment", citizens' willingness to pay for redistribution is elicited and analyzed with respect to their possible economic determinants. Based on the simple model that relates choices to the attributes of redistribution only, the average Swiss citizen is shown to require a compensation of some CHF 12 per month for an additional percentage point of GDP devoted to public redistribution. In addition, a very marked status quo bias would have to be overcome. Furthermore, several hypotheses concerning the determinants of the demand for redistribution can be tested without significant supply-side influences. By including one of three measures of current economic well-being at a time, the extended model allows us to test a static hypothesis, stating that demand for redistribution decreases with current economic well-being. However, it is found to increase with level of education and (in part) with personal income as well as higher self-positioning on a social scale, thereby contradicting the conventional Meltzer-Richard (1981) model. With the inclusion of five measures of social mobility, the dynamic Prospect of Upward Mobility (POUM) hypothesis [Hirschman and Rothschild (1973), Benabou and Ok (2001)] could be tested as well. It is shown to receive partial empirical support.

Due to the mixed empirical evidence bearing on the economic determinants, Chapter 4, entitled "Do religious beliefs explain preferences for income redistribution? Experimental evidence", focuses on behavioral factors, in particular, on religious and cultural beliefs as possible determinants of preferences for income redistribution. Estimated marginal willingness to pay (WTP) is positive among those who do not belong to a religious denomination, and negative otherwise. However, the marginal WTP is shown to increase with a higher degree of religiosity. Moreover, those who state that luck or connections rather than work effort determine economic success exhibit significantly higher WTP values than those who believe in the 'just world'. In summary it can be said that cultural and religious beliefs seem to provide far better predictions of demand for redistribution than the economic factors treated in Chapter 3.

Chapter 5, "Is the welfare state sustainable? Experimental evidence on citizens' preferences for redistribution", goes beyond the positive analysis of preferences for redistribution and contributes to the political debate by deriving implications for social policy. We argue that the sustainability of the welfare state ultimately depends on citizens' preferences for income redistribution. Estimated marginal willingness to pay (WTP) is positive among those who deem benefits too low, and negative otherwise. However, even those who state that government should reduce income inequality exhibit a negative WTP on average. The major finding is that estimated average WTP is maximum at 21% of GDP, clearly below the current value of 25%. Thus, the present Swiss welfare state does not appear sustainable.

Guidance for policy is also at the center of Chapters 6 and 7, dealing with optimal regulation instruments in markets for renewable energy. Chapter 6, "Promoting renewable electricity generation in imperfect markets: price vs. quantity policies", contrasts price and quantity control policies, starting within a framework of a perfectly competitive energy market and extending the analysis for the case of market imperfections. The search for economically efficient policy instruments designed to promote the diffusion of renewable energy technologies in liberalized markets has led to the introduction of quota-based tradable 'green' certificate (TGC) schemes for renewable electricity. However, there is a debate whether the TGC, a quantity control policy, is advantageous in comparison with guaranteed feed-in tariffs, a price control policy. In this chapter it is shown that both policies are equivalent in terms of social welfare if the electricity markets are perfectly competitive. The main result of the essay is, however, that the price control policy dominates the quantity control policy in terms of social welfare if the assumption of perfect competition is relaxed.

Chapter 7, "Renewable energy policy in the presence of innovation: does government pre-commitment matter?", extends the basic model of Chapter 6 by considering a perfectly competitive market with a possibility of technological innovations. Once again, guaranteed feed-in tariffs for electricity from renewables and tradable green certificates are contrasted from the point of view of social welfare as well as that of dynamic efficiency. The decisions about the technological innovation are modeled in a game-theoretic framework. In terms of social welfare, subsidy and quota policies are shown to be equivalent as in the static model of Chapter 6. The main finding is that subsidy policies are preferable in terms of dynamic efficiency. Further, no pre-commitment policies are shown to be at least as good as the pre-commitment ones.

Note that Peter Zweifel co-authored Chapters 2, 3, 5, and 6, Reinhard Madlener co-authored Chapters 6 and 7, Sule Akkoyunlu, Chapter 2, and Weiyu Gao, Chapter 6. Chapter 3 is currently under review at *International Tax and Public Finance*, Chapter 4 at *CESifo Economic Studies*. Chapter 6 has been invited at *Energy Economics* for a resubmission. Chapter 2 will be submitted to the *Swiss Journal of Economics and Statistics*, Chapter 5 to *Public Choice*. Chapter 7 still has a working paper status. Its abridged version will be submitted to a journal on industrial organization.

This introduction is concluded with a note to the readers concerning the structure of this dissertation. Each chapter of this dissertation can be considered as self-contained, in most cases each having its own appendix. Since Chapters 3 to 5 are based on the Discrete Choice experiment conducted in 2008, methodological sections as well as appendices in these chapters are identical. However, for the sake of better readability, all references across chapters are made explicit. Moreover, the common list of references appears at the end of this dissertation.

Why does the amount of income redistribution differ between the United States and Europe? The Janus face of Switzerland

Sule Akkoyunlu, Ilja Neustadt and Peter Zweifel

Abstract: In this paper, the amount of income redistribution in the United States, the European Union, and Switzerland is compared and empirically related to economic, political, and behavioral determinants elaborated in the literature. Lying in between the two poles, Switzerland provides unique evidence about the relative merits of competing hypotheses. It tips the balance against the economic explanation, which predicts more rather than less income redistribution in the United States compared to the EU. It only weakly supports the political model linking proportional representation and multiparty structure (which also characterize Switzerland) to redistribution; yet the Swiss share of transfers in the GDP is low. Behavioral explanations receive a good deal of support from the case of Switzerland, a country that shares with the United States the belief that hard work rather than luck, birth, connections, and corruption determine wealth. In this way, the Janus face of Switzerland may help to explain the difference in the amount of U.S. and EU income redistribution.

Keywords: income redistribution, income mobility, openness, proportional representation, beliefs, religiosity

JEL classification: D31; D63; H53; I31

Chapter 2

Why does the amount of income redistribution differ between the United States and Europe? The Janus face of Switzerland

2.1 Introduction

The objective of this paper is to compare the amount of income redistribution of the United States, the European Union (EU), and Switzerland. While a European country, Switzerland is not a member of the EU and has some institutional features reminiscent of the United States. Like the Roman god Janus it is therefore predicted to "look both ways". Indeed, EU social programs will be found to be more extensive, generous, and pro-poor and tax systems to be more progressive than those of United States. Invariably, Switzerland stands in between. What are the economic, political, and behavioral factors that may be responsible for this? Possible economic explanations are the variance and skewness of the before-tax income distribution, the social costs of taxation, expected future changes in income for median voters and volatility of income over time. However, Alesina et al. (2001) argue that these economic determinants cannot explain observed differences in redistributive policies between the United States and the EU. They find that while the before-tax income in the

United States has higher variance and more skewness than in the EU, redistribution in the United States is less although the deadweight losses from taxation seem to be about the same. Switzerland will be shown to lie in between. On the other hand, the "Prospect of Upward Mobility" (POUM) hypothesis, originally suggested by Hirschman and Rothschild (1973) and formulated by Benabou and Ok (2001), is confirmed by Alesina and La Ferrara (2005) who empirically show that people with high expected future income do not favor redistribution in the United States[1]. In the EU, income mobility is relatively low, with Switzerland again situating in between but having a lower level of public expenditure than the EU average. Thus, contrary to theoretical predictions, we observe a negative correlation between income mobility and public expenditure on the aggregate level. As to political explanations, Alesina and Glaeser (2004) cite U.S. institutions that prevent minorities from gaining political power which could be used for income redistribution. At the federal level, the United States applies majority rule for election to the Congress and for president; moreover, courts have consistently been rejecting popular attempts at redistribution. The constitutions of EU member countries are more oriented toward proportional representation and less toward protection of private property. Switzerland on the one hand has a degree of proportional representation that even exceeds the EU average; on the other hand, its courts strongly protect private property [cf. Moser (1994)]. Extensive direct democratic control might serve to limit public welfare spending while enforcing efficiency in redistribution. But then, Switzerland is comparable to the EU average when it comes to the amount of transfers and subsidies. The behavioral explanations for redistribution [Fong et al. (2006)] emphasize reciprocal altruism. This hypothesis states that U.S. voters dislike giving money to the poor whom they perceive as lazy. Moreover, Gilens (1999) and Alesina and Glaeser (2004) argue that troubled race relations are a major reason for the absence of an American welfare state. EU citizens, by way of contrast, tend to believe that the poor have been unfortunate, and until recently, immigration from non-white countries has been too limited to make race a relevant category. Between these two poles, Switzerland seems to be similar to the United States in all of the three dimensions cited above, giving rise to the correct prediction that it spends relatively little on public welfare. It should be noted that all of these explanations abstract from the

[1] Using a data set from Russia, Ravallion and Lokshin (2000) have shown that even those who are currently rich may tend to support redistribution if they expect their welfare to fall. This is known as a "tunnel effect". Molnár and Kapitány (2006a,b) find that people who have no clear knowledge about the immediate and the distant future favor redistribution more than those with negative expectations.

incentives of politicians, acting as entrepreneurs, to redistribute income and wealth in order to secure (re)election [cf. Brunner and Meckling (1977), Cukierman and Meltzer (1986)]. This ultra-political explanation hinges on the fact that the cost of redistribution usually takes on the form of efficiency losses that have to be borne by all citizens, whereas its benefits can be channeled to those lobbies that provide support or those voters who are pivotal at the next election. Of course, the institutional differences cited above make it easier for politicians to pursue their objectives in some countries and more difficult in others. Yet, politicians have a permanent incentive to push back those constraints that limit their freedom of action. In all, this hypothesis predicts that redistribution occurs largely regardless of preferences in the population. For simplicity, it will not be pursued in detail but may serve as an explanation of why the amount of redistribution keeps growing over time [for an analysis in the case of social health insurance, see Zweifel (2007)]. The structure of this paper is as follows. In Section 2.2, the size and structure of redistribution in the United States, selected EU countries, and Switzerland are presented. Section 2.3 tests the economic explanations for redistribution, which are contradicted by the case of Switzerland. Section 2.4 again finds that political explanations are not confirmed by Swiss experience. Section 2.5 presents behavioral determinants which are not only successful in explaining the differences between the United States and the EU but are also confirmed by Switzerland. Section 2.6 provides a final assessment of the determinants of public social spending in a multivariate model. Section 2.7 concludes.

2.2 Size and structure of redistribution in the United States, the EU, and Switzerland

In this section, the basic facts concerning redistribution in the United States, the EU, and Switzerland are presented, starting first with government spending and revenue, and then turning to regulation designed to achieve income redistribution, such as minimum wage laws.

2.2.1 Government spending

Table 2.1 shows the size and composition of government expenditure. Total government expenditure in the EU-15 averages 46 percent of GDP; it reaches 53 percent in France and even 56 percent in Sweden but only 37 percent in the United States. Switzerland is just below the

country	total government expenditure	consumption (appropriation account)			subsidies	social benefits	fixed investment
		total consumption	*goods and services*	*wages*			
US	36.6	15.8	6.1	9.7	0.4	12.0	3.3
EU-15	46.0	20.4	10.2	10.2	1.2	16.3	2.5
Austria	49.2	18.0	8.7	9.3	3.1	18.3	1.1
France	53.4	23.6	10.5	13.1	1.5	17.8	3.4
Germany	45.7	18.3	11.1	7.2	1.2	18.6	1.4
Sweden	55.5	26.8	11.1	15.7	1.6	16.7	3.1
UK	45.0	22.0	10.6	11.4	0.4	13.0	1.8
Switzerland	36.3	10.9	2.8	8.1	4.0	12.0	2.3

Source: authors' calculations based on data from OECD Economic Outlook database (No 82, Dec. 2007). Details may not sum to totals because of excluded categories.

Table 2.1: Composition of general government expenditure in percent of GDP, 2006

U.S. value with 36 percent. However, it is the share of transfers (subsidies and social benefits) where differences are most marked. In fact, the sum of these categories amounts to 17.5 percent of GDP in the EU compared to 12.4 percent in the United States. Here Switzerland sides with the EU, its share being 16 percent.

Table 2.2 presents the breakdown of social expenditure (which notably includes old-age benefits). First, the United States is far below the EU average with 15 and 24 percent of GDP, respectively. Switzerland even exceeds the EU average with 26 percent, coming close to full-fledged welfare states such as Germany (27 percent) and France (29 percent). The main reason are old-age benefits, where U.S. public expenditure makes up a low 5 percent of GDP, compared to the EU share of 9 and the Swiss share of 12, respectively. In relative terms, the differences in family benefits are even more pronounced. Here, the United States spends one-fifth of the EU value (0.4 compared to 2.2 percent of GDP in the EU-15), with Switzerland once more falling in between (1.2 percent). However, this does not necessarily mean that countries such as France, Germany, or Sweden are pro-poor because social security systems typically redistribute from the young to the old.

2.2.2 Government revenue

Government expenditure of a country may be pro-poor; yet if it is financed in a highly regressive manner, the net effect of government activity may turn pro-rich. Table 2.3 summarizes the composition of government revenue in the EU, the United States, and Switzerland. First of all, the EU governments claim a much larger share of the GDP (46 percent on average)

country	total	old-age	family	unemployment	health	incapacity	other
US	14.7	5.3	0.4	0.5	6.2	1.1	1.2
EU-15	23.8	8.8	2.2	2.1	6.1	2.9	1.7
Austria	26.0	10.7	2.9	1.3	5.2	2.5	3.4
France	28.5	10.6	2.8	2.9	7.2	2.1	2.9
Germany	27.4	11.7	1.9	2.3	8.0	2.3	1.2
Sweden	29.8	9.2	3.8	2.4	7.4	5.2	1.8
UK	21.8	8.1	2.2	0.6	6.1	2.5	2.3
Switzerland	26.4	11.8	1.2	1.0	6.4	3.8	2.2

Source: OECD (2004) Social Expenditure database.

Table 2.2: Public social expenditure in percent of GDP, 2001

country	total receipts	tax revenue				social security contributions	property income	other
		direct taxes			indirect taxes			
		total	*households*	*businesses*				
US	34.0	13.6	10.3	3.3	7.3	7.0	0.8	5.3
EU-15	45.6	12.2	9.3	2.9	13.6	15.5	0.9	3.4
Austria	47.8	13.1	10.7	2.4	14.0	16.0	1.2	3.5
France	50.8	11.8	8.7	3.1	15.4	18.3	0.7	4.6
Germany	43.8	10.6	9.2	1.4	12.1	17.3	0.6	3.2
Sweden	57.9	20.2	16.5	3.7	17.1	13.2	2.2	5.2
UK	41.9	17.2	13.1	4.1	12.8	8.4	0.6	2.9
Switzerland	35.4	14.9	11.3	3.6	7.2	7.1	1.4	4.8

Source: authors' calculations based on data from OECD Economic Outlook database (No 82, Dec. 2007).

Table 2.3: Composition of general government revenue in percent of GDP, 2006

than their U.S. counterpart (34 percent). The figures do not match precisely those of Table 2.1 because in 2006, governments were accumulating debt at a different pace. The Swiss government showed the best budgetary discipline among the countries sampled, its expenditure share in the GDP of 36 percent exceeding its revenue share of 35 percent by relatively little. Second, governments substitute direct taxes by social security contributions. In the United States, the ratio of the former to the latter is 14/7, while in the EU it amounts to 12/16, and in Germany, even 11/17. With a ratio of 15/7, Switzerland definitely resembles the United States here. Thus, in terms of direct taxation, some EU countries might look like tax havens compared to the United States and Switzerland but they make up by charging much higher social security contributions. Whether this reflects a more marked pro-poor orientation depends on the relative progressiveness of social security and income taxation. By way of contrast, indirect taxation is generally regarded as regressive. The ratio of direct to indirect taxes is 14/7 for the United States, 12/14 for the EU, but 15/7 for Switzerland. On

country	1980	1990	2000	2006
US	12.9	13.7	14.9	16.6
EU-15	21.3	21.8	21.7	21.7
Austria	25.5	26.6	27.4	26.1
France	21.3	22.7	24.2	25.4
Germany	21.5	21.2	24.0	23.2
Sweden	25.9	27.8	25.9	25.7
UK	16.3	15.1	17.0	18.4
Switzerland	n.a.	14.9	18.9	20.5

Source: authors' calculations based on data from OECD Economic outlook database (No. 82, Dec. 2007).
Table 2.4: Government expenditure on subsidies, social benefits and other current transfers in percent of GDP, 1980-2006

this account, both the United States and Switzerland look more pro-poor than the EU, with France (12/15) marking an extreme.

However, not only does the status quo reveal important differences; developments during the last few decades differ, too. Table 2.4 tracks the government expenditure categories "subsidies" and "social benefits" of Table 2.1 (complemented by "current transfers", not evidenced there) since 1980. By that time, countries such as Austria, France, and Germany were full-fledged welfare states with GDP shares above 20 percent, while the United States stood at 13 percent. Since then, it has caught up somewhat, reaching some 17 percent in 2006. While data for 1980 are not available for Switzerland, in 1990 its share of 15 percent was close to that of the United States. However, Swiss transfer payments have increased particularly fast since then, attaining 21 percent in 2006, not far from the EU average of 22 percent anymore. Summing up the findings so far, Switzerland resembles the EU in terms of its government expenditure but is more similar to the United States in terms of its government revenues. It used to be close to the United States with regard to transfers but has been approaching the EU during the last two decades.

2.2.3 Redistribution through private charity

The preceding data suggest that EU countries and Switzerland provide more public welfare than the United States. However, the World Values Survey (Table 2.5) shows that Americans engage in more private provision of welfare through charity than EU and Swiss citizens. Roberts (1984) hypothesizes that public provision of welfare in part crowds out private charity.

country	active member	inactive member	not a member
US (1995)	27.3	14.9	57.8
Germany (1997)	7.9	13.8	78.3
Sweden (1996)	6.7	15.8	77.5
Switzerland (1996)	5.8	15.3	78.9

Source: World Values Survey.
Table 2.5: Membership in charitable organizations in percent of total population, 1995-1997

country	percent of GDP
US	1.67
France	0.14
Germany	0.22
Ireland	0.47
Netherlands	0.45
UK	0.73
Switzerland	0.37

Sources: Charities Aid Foundation, ZEWO Foundation.
Table 2.6: Charitable giving as a share of GDP, 2005

Potential donors, seeing government transfers on the rise, have a weakened motivation to give. Being altruistic, they might also be willing to donate through the government. However, the symmetry of substitution effects leads to the prediction that those who donate privately prefer to limit public transfers.

Therefore, a low level of public expenditure in the United States could be partially explained with high private donations. Table 2.5 tends to support this view. In the United States, 27 percent of the population report to actively participate in a charitable organization, compared to 8 percent in Germany and a mere 6 percent in Switzerland. Conversely, only 58 percent of U.S. citizens indicate not to be involved in any charitable organization, whereas their European counterparts are close to the 80 percent level. Table 2.6 shows that the amount of charity giving in the US is also higher than in EU countries and Switzerland, suggesting that public transfers cause a reduction in voluntary donations in Europe, as predicted by the crowding-out literature. On this score, Switzerland definitely sides with the EU rather than the United States.

country	1960	1970	1980	1990	2000	2005
US	42.3	39.3	39.7	42.7	45.7	45.0
EU-15	35.1	35.1	31.2	29.6	30.3	29.9
Austria	-	29.5	31.6	26.3	29.2	26.0
France	49.0	39.8	36.4	28.0	28.2	28.0
Germany	38.0	39.2	36.6	30.8	29.8	28.0
Sweden	-	29.5	19.4	21.9	27.2	23.0
UK	25.5	25.4	25.3	33.5	34.6	35.0
Switzerland	-	-	35.9	33.8	31.8	31.1

Source: WIID database, World Institute for Development Economics Research 2006.

Table 2.7: Gini coefficients in the U.S., EU, and Switzerland, 1960-2005

2.3 Economic explanations of income redistribution

One of the main economic explanations of income redistribution states that the more marked the pre-tax income inequality, the higher the demand and the political pressure for redistribution. This is the basic idea behind the Romer-Roberts-Meltzer-Richard (RRMR) model[2] stating that the lower the income of the median voter relative to the income of the average voter, the higher the level of taxation and redistribution.

Indeed, U.S. income inequality was high in 1960 [Gini coefficient of 42, see Deininger and Squire (1996)] and has been again increasing since 1970 to reach a Gini of 45 in 2005. In the same period, the average value of EU countries has fallen from 35 to 30. The most notable decrease occurred in France, from 49 to 28. As to Switzerland, the first measurement dates back to 1980. Since then, its Gini coefficient has been decreasing even faster, from 36 to 31 (the U.S. and EU values being 43 and 30, respectively, at the time). Therefore, in 1980 Switzerland lay right in between the two poles but has been approaching the EU fast since. In view of the marked pre-tax income inequality in the United States, combined with low government expenditure and few labor market interventions, the RRMR model finds very weak support by the evidence. Alesina and Giuliano (2009) point out that the main failure of this model rests on its simplistic assumptions, viz. the ‘one person, one vote’ rule and the median-voter outcome. Barenboim and Karabarbounis (2008) show empirically that the very rich have more weight above and beyond the ‘one person, one vote’ rule in the political process, while the very poor do not vote at all. Neustadt and Zweifel (2009) [see Chapter 3 of this dissertation] conduct a discrete choice experiment in Switzerland and elicit willingness

[2]Romer (1975); Roberts (1977); Meltzer and Richard (1981)

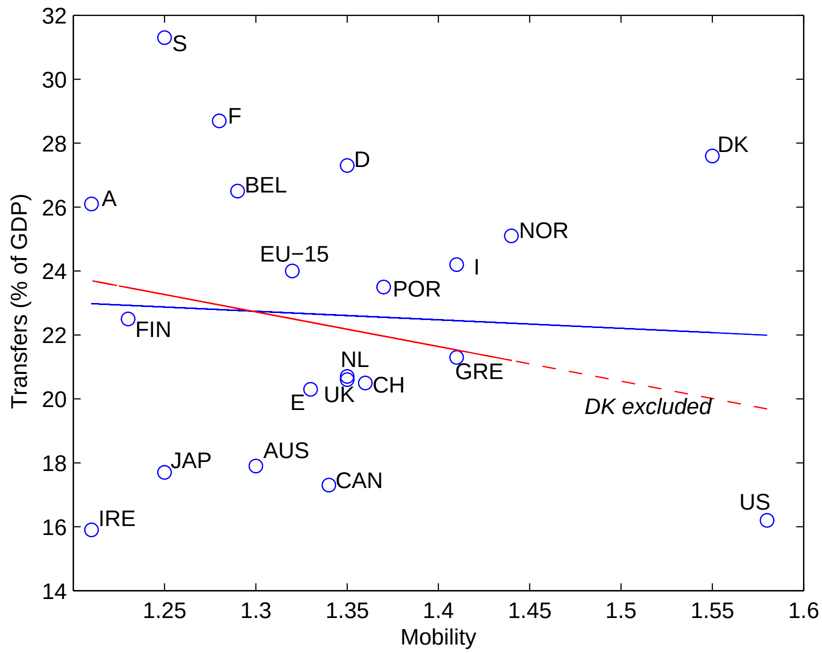


Figure 2.1: Relationship $y = \hat{\alpha} + \hat{\beta}x$ between transfers y (% of GDP in 2007) and income mobility parameter x defined as the ratio of the income in the fourth quintile to the income in the third quintile

$\hat{\alpha} = 26.2112$, $\hat{\beta} = -2.6698$, $\bar{R}^2 = 0.0036$ for the whole sample (the t statistic is -0.2636, i.e. not significant); $\hat{\alpha} = 36.81$, $\hat{\beta} = -10.838$, $\bar{R}^2 = 0.0493$ if Denmark excluded (the t statistic is -0.966, i.e. still not significant). Country labels: A=Austria, AUS=Australia, BEL=Belgium, CAN=Canada, CH=Switzerland, D=Germany, DK=Denmark, E=Spain, F=France, FIN=Finland, GRE=Greece, I=Italy, IRE=Ireland, JAP=Japan, NL=Netherlands, NOR=Norway, POR=Portugal, S=Sweden, UK=United Kingdom, US=United States, EU-15=simple average of old EU member countries without Luxembourg (A, BEL, D, DK, E, F, FIN, GRE, I, IRE, NL, POR, S, UK). Data source: CIA World Factbook 2008.

to pay for income redistribution. Their analysis of preference heterogeneity with respect to current economic well-being shows that willingness to pay for redistribution increases with income and education, contradicting the RRMR model.

However, as hypothesized by Benabou and Ok (2001), earnings mobility may dampen a poor but forward-looking voter's enthusiasm for redistribution [for empirical support using U.S. data, see Alesina and La Ferrara (2005)]. In their study of willingness to pay for redistribution, Neustadt and Zweifel (2009) [see Chapter 3] use five alternative mobility measures and show that this Prospect of Upward Mobility (POUM) hypothesis receives partial empirical support, albeit for only four of five measures used. However, individuals with no mobility at all display the highest resistance against redistribution, contradicting the POUM hypothesis but underscoring the importance of a high status quo bias. As a partial test, Figure 2.1 plots public transfers (GDP share) against the ratio of average income in the (relatively wealthy) fourth and average income in the (middle class) third quintile. Admittedly, this is a rather poor measure of mobility, as discussed in Muren and Nyberg (2005). One would prefer to take into account probabilities of transition from the third to the fourth quintile. However, the data on these transition probabilities are currently available for six countries only [see OECD (1996)]. Still, since the quintile transition probabilities are shown to be quite similar among the OECD countries [Muren and Nyberg (2005)], a large inter-quintile income difference can serve as a rough indicator of income mobility. In the United States, the difference between the third and the fourth quintile is indeed large (1.55 or 55 percent more income), whereas it is around 1.3 in the EU on average. When the outlier Denmark (DK in Figure 2.1) is excluded as an outlier to a negative relationship, the negative slope of the regression becomes slightly more marked, providing weak support for the POUM hypothesis of Benabou and Ok (2001). However, the coefficient of determination remains low, and United States (US in Figure 2.1) as well as Canada (CAN), Australia (AUS), Japan (JAP), and Ireland (IRE) lie far below the regression line.

Some authors establish a link between openness of the economy and the level of income redistribution by postulating the compensation hypothesis [Cameron (1978); Katzenstein (1985); Garret (2000); Adsera and Boix (2002)]. This hypothesis states that small open economies compensate their losers from international liberalization with government interventions in the domestic economy, mainly with an increase in transfer payments. Higher levels of trade imply growing risks associated with the international business cycle and thereby cause higher levels of income volatility and income inequality. As stated in Section 2.1, under the veil of uncertainty, risk-averse individuals may be willing to support income redistribution programs, especially if designed to help those who suffered an unexpected loss in their assets (health, wealth, wisdom, i.e. skills). Emphasizing the former effect, viz. that open economies expose citizens to more income volatility because they are subject to external shocks, Rodrik (1998) relates income redistribution to the openness of the economy. Other authors [Adsera and Boix (2002), Balcells Ventura (2006)] emphasize the latter effect, the increasing inequality based

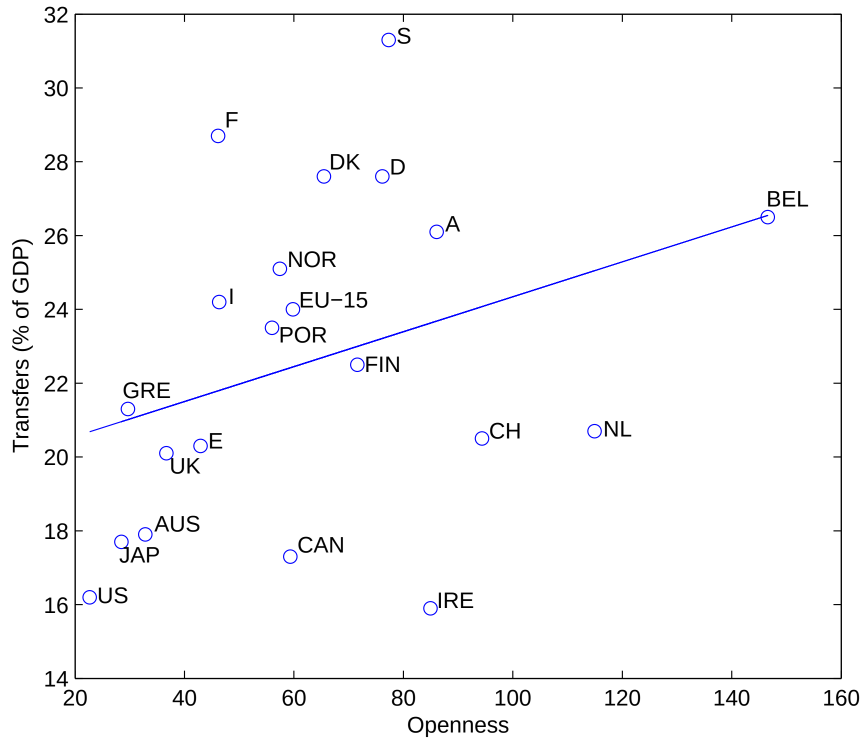


Figure 2.2: Relationship $y = \hat{\alpha} + \hat{\beta}x$ between transfers y (% of GDP in 2003) and openness x, defined as sum of exports and imports 2007 over GDP in 2007

$\hat{\alpha} = 19.609$, $\hat{\beta} = 0.0473$, $\bar{R}^2 = 0.1094$, t statistic is 1.433 (not significant). Data source: CIA World Factbook 2008.

on the idea that openness to trade creates winners and losers within economies. They show that the impact of openness on income redistribution crucially depends on income per capita and the size of potential loser sectors. While trade has a positive effect on the size of the public sector in rich countries (those abundant in high-income factors), it negatively affects the level of income redistribution in poor countries. Figure 2.2 plots[3] transfers as a share of GDP against an indicator of openness, the ratio between the sum of exports and imports and GDP. Indeed, the United States, being a rather closed economy, has the lowest transfer share. And in general, increased openness does go along with more transfers for 'rich' OECD

[3]Given that transfers are associated with inefficiencies, one could argue that transfers as a 'type of insurance against the vagaries of openness' should progressively increase with openness. However, a regression of $transfers$ on *openness* and $(openness)^2$ yields a negative but insignificant term. The nonlinear relationship between political (and social) openness and welfare is examined by Koster (2008). The author finds weak evidence of nonlinearity for social openness, but no evidence for political openness.

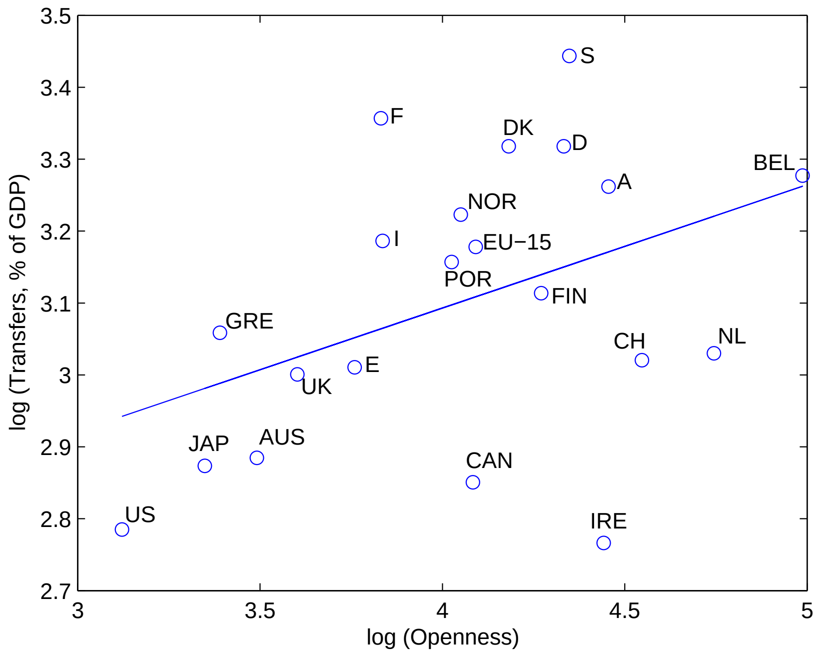


Figure 2.3: Relationship $y = \hat{\alpha} + \hat{\beta}x$ between the natural logarithm of transfers y (in % of GDP in 2003) and the natural logarithm of openness x, defined as ratio of the sum of exports and imports 2007 over GDP in 2007

$\hat{\alpha} = 2.4068$, $\hat{\beta} = 0.1716$, $\bar{R}^2 = 0.1716$, t statistic is 1.977 (significant at the 6.4 percent level). Data source: CIA World Factbook 2008.

countries, thus seemingly supporting the result of Balcells Ventura (2006). However, with a t statistic of 1.433, this bivariate regression does not provide conclusive evidence of a positive relationship.

In a next step, Figure 2.3 plots the natural logarithm of the share of transfers in GDP against the natural logarithm of the indicator of openness as defined above. Now the t statistic has the value of 1.977 and thus implies weak evidence (at the 6.36 percent significance level) of a positive elasticity of transfers as a share of GDP with respect to the indicator of openness. One might argue that openness as defined by Rodrik (1998) fails to measure the impact of foreign trade shocks on the welfare of a population. Shifts in the terms of trade, however, directly indicate changes in the gains from trade a country can reap and hence in welfare. During the period 1960-2006, the U.S. terms of trade exhibited a standard deviation of 0.133 percentage points p.a. While comparable data are lacking for the EU, Austria and Germany come in with 0.05 and 0.085 points, respectively[4]. Once more, Switzerland is in between with 0.106 points. Note that the high U.S. value would lead one to predict a high amount of redistribution, contrary to the empirical evidence. At best, one could argue that social mobility in the United States serves as a substitute for redistributive policies.

On the whole, economic explanations do not seem to be very successful in predicting the amount of income redistribution, at least when relying on government expenditure and transfers as indicators. If one is willing to use Janus-faced Switzerland as a test case, this country is never even close to the regression line. Thus, it causes the balance to be tipped against economic explanations.

2.4 Political explanations

The United States, the EU, and Switzerland differ in terms of their political institutions. The first aspect relates to the electoral level. The United States has a majoritarian system where the plurality rule is applied in federal elections (i.e. each district delegates the representative with the most votes), while all EU countries (with the exception of the United Kingdom and France) have proportional representation. Proportional representation tends to produce multiparty parliaments and governments, while majority rule favors a strict two-party system as in the United States or a multiparty system dominated by two players as in the United Kingdom. The political science literature [Lizzeri and Persico (2001), Milesi-Ferretti et al. (2002), Persson and Tabellini (2000, 2003)] predicts that proportional representation tends towards universal programs benefitting various groups (pensioners, workers, poor, minorities, etc.), while majority rule results in targeted 'pork barrel' programs. Persson and Tabellini (2000, 2003) find supporting empirical evidence in that countries with proportional representation have GDP share of government expenditure that ceteris paribus is 5 percentage points higher than with majority rule.

Figure 2.4 illustrates the effect of electoral rules on fiscal policy, plotting transfers as a share of GDP against a measure of proportional representation for most OECD countries. There is indeed weak evidence of a positive correlation. While the EU-15 is close to the regression line, the United States constitute an outlier. This is true of Switzerland too, in spite of its

[4] Authors' calculations from OECD Economic Outlook Database, No. 82, Dec. 2007, World Bank and WMM.

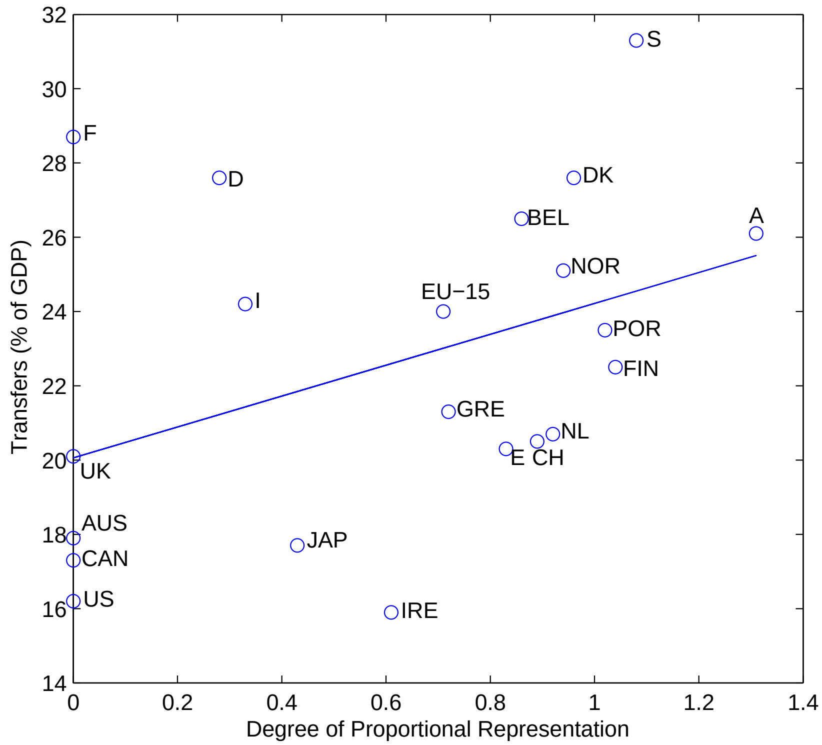


Figure 2.4: Relationship $y = \hat{\alpha} + \hat{\beta}x$ between transfers y (in % of GDP in 2003) and the degree of proportional representation x

$\hat{\alpha} = 20.058$, $\hat{\beta} = 4.16$, $\bar{R}^2 = 0.1663$, t statistic is 1.91 (significant at the 7.22 percent level). Degree of proportional representation is the natural logarithm of the size of electoral districts, defined as the number of electoral districts in a country divided by the number of seats in the lower or single house for the most recent legislature. Data source: CIA World Factbook 2008, Persson and Tabellini (2003).

system of proportional representation and a system with several strong parties that is similar to continental EU countries. The reasons for this divergence are discussed below.

The second aspect of political institutions relates to the government. The United States has a presidential system while all EU countries are parliamentary democracies (with the exception of France, whose government is controlled by the majority in the parliament, however). Presidential regimes at first sight result in a concentration of power; however, they tend to have a stronger separation of powers designed to prevent abuse [cf. Persson et al. (1997)]. Since this abuse goes along with increased government expenditure and transfers to supporting clientele groups, presidential systems are predicted to induce less income redistribution.

The third aspect is the party system. Barriers to entry for parties are particularly high in the United States, likely due to the country's vast size and low population density, both of which help to diffuse social conflict. This has resulted in the absence of a strong socialist party, whereas the European left was able to organize and divulge its ideas, resulting in a higher amount of income redistribution.

The fourth aspect of political institutions of relevance for redistribution is fiscal decentralization. This creates obstacles to an excessive role for the central government in fiscal matters, making it more difficult to tax the rich in some part of the country in favor of the poor localized in other parts. Again, the United States is characterized by a higher degree of fiscal federalism than most EU countries [Inman and Rubinfeld (1992)], which may help explain its lower amount of income redistribution.

As to Switzerland, it is on the U.S. side on items three and four (Supreme Court, fiscal decentralization) but on the EU side on items one and two (proportional representation, low barriers to entry for political parties). However, the distinguishing feature of Switzerland in this context is its direct democracy with popular initiatives and referenda. Feld et al. (2007) find that public expenditure tends to be better tailored to the needs of the electorate in direct than in representative democracies. If the electorate wishes to be pro-poor, Swiss redistributive policies might attain its objectives at a lower value of total transfers than representative democracies. As noted in the context of the first aspect cited (proportional vs. majoritarian representation), this observation is not discriminating because Switzerland is below the regression line in Figure 2.4. However, the Netherlands and Spain, two countries with almost no direct democratic control, have the same GDP share of transfers as Switzerland. Therefore, direct democratic control cannot alone explain why Switzerland has low transfers in spite of its high degree of proportional representation.

Summing up, four aspects of political institutions seem to be relevant for income redistribution. One of them (degree of proportional representation) could be quantified; it did show the predicted relationship with the transfer share in GDP. Using again Switzerland for corroborating evidence, the country shares institutional features both with the United States and the EU. However, it is unique in its degree of direct democratic control, yet has the same GDP share of transfers as the Netherlands and Spain, two countries with quite different political institutions. Therefore, political explanations appear only slightly more convincing than the economic ones.

2.5 Behavioral explanations

Behavioral explanations of income redistribution importantly revolve around the concept of imperfect altruism. While prefect altruism is exclusively governed by recipients' preferences, imperfect altruism also reflects donor preferences. In particular, it predicts that people will oppose public welfare if they believe that recipients take advantage of the system, a behavior that is often attributed to members of ethnic minorities. Alesina et al. (1999), Alesina et al. (2001), Luttmer (2001), Alesina and Glaeser (2004), and Luttmer and Singhal (2008) find that people oppose redistribution favoring ethnic or racial groups other than their own as well as minorities that are overrepresented among the poor.

As a first piece of evidence, Figure 2.5 plots the bivariate relationship between public transfers and ethno-linguistic fragmentation. While most EU countries are quite homogeneous with respect to ethnicity and language, Belgium and Spain display a degree of heterogeneity that exceeds that of the United States (Canada is the extreme case here). There is a negative correlation, supporting the hypothesis. Switzerland has a high heterogeneity too, reflecting the strong division between the German-speaking, French-speaking, and Italian-speaking parts of the country. However, this time it lies right on the regression line, providing corroborating evidence.

A second aspect of fragmentation is migration. Immigration serves to increase the heterogeneity of a society. The net migration rate is defined as the difference between the immigration rate and the emigration rate. It would be preferable to consider the immigration rather than the net migration rate. However, data on the immigration rate are available for selected countries only. Still, in all OECD countries under consideration the rate of emigration is significantly lower than the immigration rate. Therefore, the net migration rate can be used as a rough approximation of the immigration rate. As Figure 2.6 shows, countries with higher net migration rates tend to spend smaller fractions of their GDP on transfers. The corresponding bivariate regression comes very close to conventional significance levels. The United States constitutes an outlier with especially low transfers, presumably due to a third aspect, racial heterogeneity (which is more pronounced there than in the majority of EU countries). Indeed,

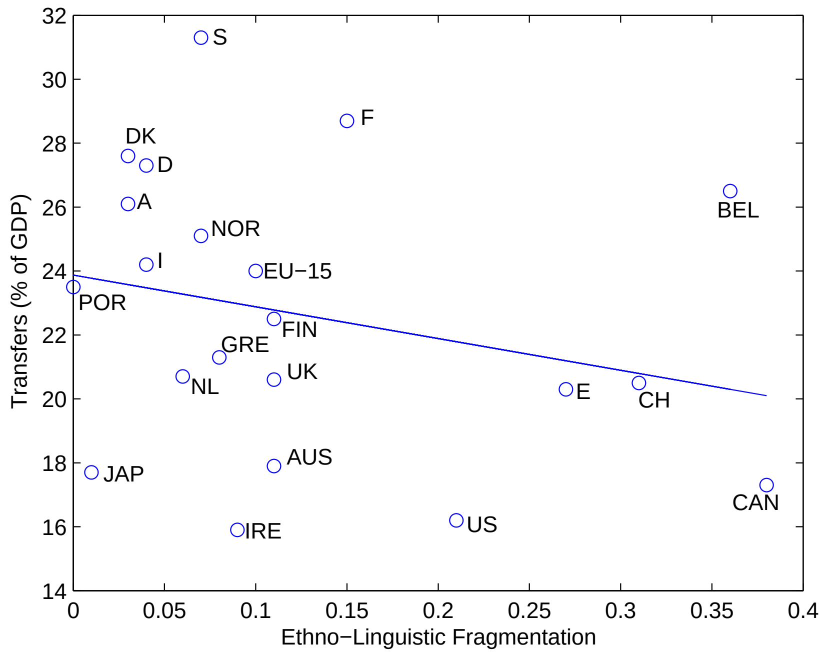


Figure 2.5: Relationship $y = \hat{\alpha} + \hat{\beta}x$ between transfers y (in % of GDP in 2003) and the ethno-linguistic fragmentation x

$\hat{\alpha} = 23.87$, $\hat{\beta} = -9.915$, $\bar{R}^2 = 0.0675$, t statistic is -1.18 (not significant). The index of ethno-linguistic fragmentation is the level of lack of ethnic and linguistic cohesion within a country, ranging from 0 (homogeneous) to 1 (strongly fragmented) and averaging five different indices, see Persson and Tabellini (2003). Data source: CIA World Factbook 2008, Persson and Tabellini (2003).

work by Kinder and Sanders (1996) reveals that racial resentment is the most powerful determinant of whites' (who are overrepresented among payers) opinions on welfare, affirmative action, school desegregation, and the plight of the inner city. Switzerland lies close to the regression line. On the one hand, its rate of net migration and share of foreign population are very high, similar to those of the United States. But on the other hand, being foreign is not necessarily associated with (permanent) poverty, similar to most EU countries.

Following Razin and Sadka (1995), the birth rate may be seen as a third indicator of fragmentation. A high rate of fertility calls for a great deal of intra-family redistribution, which squeezes out public transfers. This argument suggests a negative correlation; however, a positive relationship cannot be excluded due to reverse causality. A high birth rate could be argued to trigger a great deal of transfers in the guise of family allowances. Moreover, many govern-

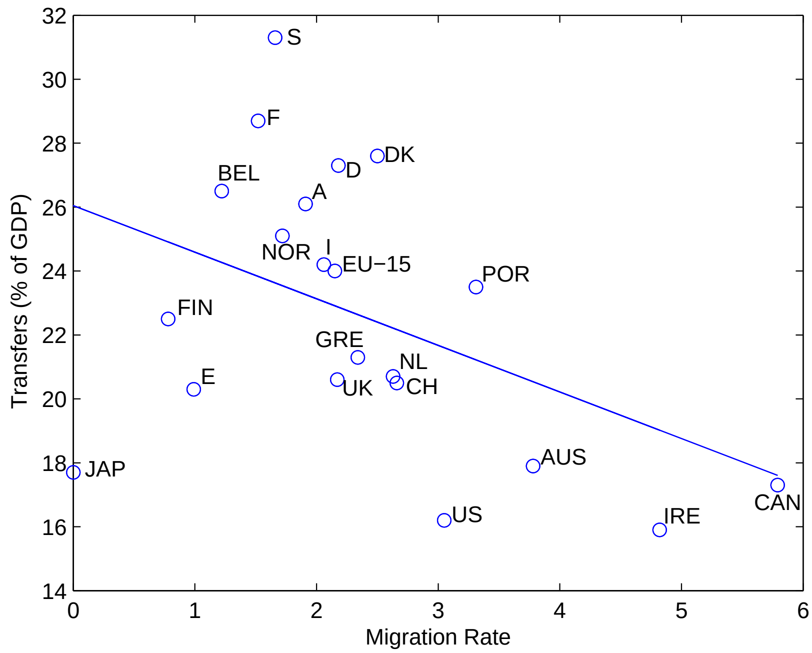


Figure 2.6: Relationship $y = \hat{\alpha} + \hat{\beta}x$ between transfers y (in % of GDP in 2003) and net migration rate x

$\hat{\alpha} = 26.047$, $\hat{\beta} = -1.458$, $\bar{R}^2 = 0.1948$, t statistic is -2.093 (significant at the 5.08 level). Data source: CIA World Factbook 2008.

ments see family allowances as a means to increase the birth rate. When transfers are plotted against the birth rate, a negative relationship obtains (see Figure 2.7). The United States has a fertility rate that is only exceeded by Ireland, one-half higher than the EU average, which reflects very low rates in countries such as Germany, Austria, and Italy. Switzerland again lies close enough to the regression line to provide some support to the hypothesis.

A fourth behavioral element is beliefs. The hypothesis is that a society who believes that luck, birth, connections, and corruption determine wealth will choose a high degree of redistribution, financed by high taxes, see Alesina and Glaeser (2004) and Alesina and Angeletos (2005). By way of contrast, the conviction that high income and wealth are the result of work effort goes along with little income redistribution. Beliefs do differ sharply between the United States and the EU. Most Americans believe that anyone can get out of poverty by hard work and that the poor remain poor only because they refuse to make the effort. By way of contrast, Europeans generally think that poverty is due to bad luck and not the in-

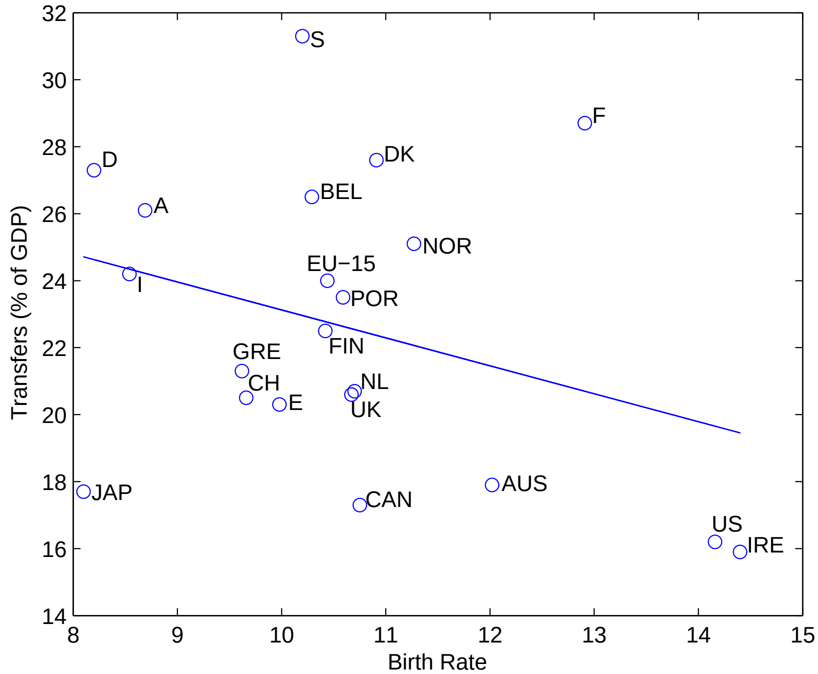


Figure 2.7: Relationship $y = \hat{\alpha} + \hat{\beta}x$ between transfers y (in % of GDP in 2003) and birth rate x

$\hat{\alpha} = 31.469$, $\hat{\beta} = -0.834$, $\bar{R}^2 = 0.1058$, t statistic is -1.526 (not significant). Data source: OECD, CIA World Factbook 2008.

dividual's responsibility. Fong and Oberholzer-Gee (2007) measure the willingness-to-pay for justice in the United States using dictator games. Dictators were given \$10 to split between themselves and recipients. The authors find that one third of the dictators are willing to pay one dollar out of ten for obtaining the information whether poverty was due to disability or substance abuse. Finally, Alesina and Giuliano (2009) show that a history of misfortune in the recent past such as unemployment and personal trauma makes people more risk-averse and less optimistic about upward mobility. These changes in beliefs are found to have a positive and significant effect on redistribution.

Figure 2.8 plots transfers against a score that ranges from 1 (hard work always brings a better life) to 10 (hard work does not bring any success). The Unites States is the observation closest to the score of 1 but still lies below the regression line. Germany (D) and Denmark (DK) mark the other extreme. With a coefficient of determination of 0.3 and a t statistic of 2.778, this is one of two best-fitting bivariate regressions designed to explain the share of transfers

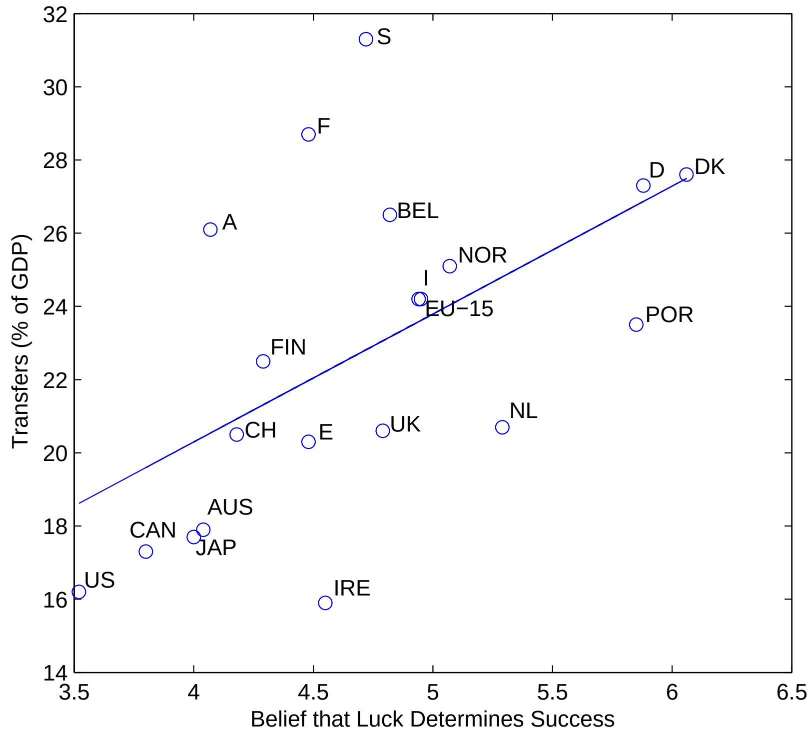


Figure 2.8: Relationship $y = \hat{\alpha} + \hat{\beta}x$ between transfers y (in % of GDP in 2003) and the belief x that luck determines success (median value for each country, measured as an index from 1 to 10, with 10 indicating strongest belief)

$\hat{\alpha} = 6.319$, $\hat{\beta} = 3.495$, $\bar{R}^2 = 0.3006$, t statistic is 2.778 (significant). Data Source: OECD, World Values Survey.

in GDP. Here again, Switzerland lies right near the regression line, lending additional support to the hypothesis.

A fifth behavioral element is political attitudes. For a long time, political scientists have been relating left-wing orientation to attitudes in favor of income redistribution [Downs (1957)][5]. However, the relationship between the political orientation of the median voter and the actual amount of redistribution (measured by the share of GDP devoted to transfers, as before) turns out amazingly weak. In Figure 2.9, political attitudes of the median voters range on a scale between 1 (left-wing) and 10 (right-wing). Note that there is little variation, with the EU-15 at 5.3 and the United States at 5.8. Switzerland lies close enough to the regression line to

[5] Frohlich and Boschmann (1986) provide supporting empirical evidence for the United States and Canada.

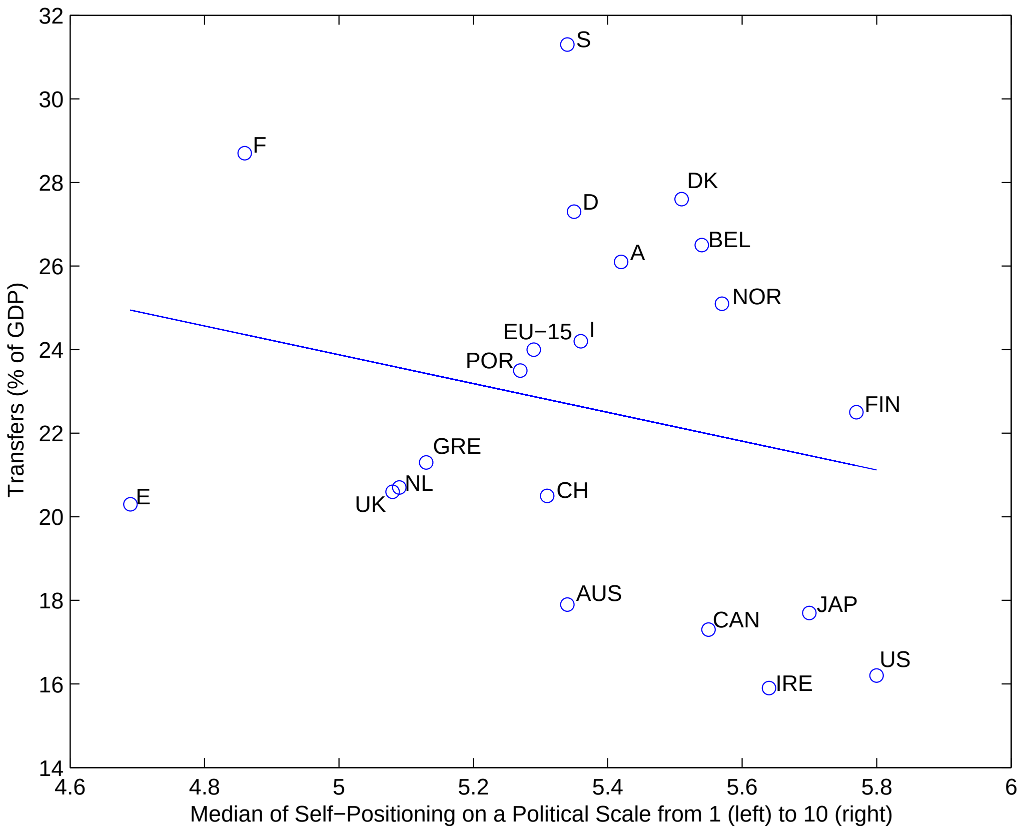


Figure 2.9: Relationship $y = \hat{\alpha} + \hat{\beta}x$ between transfers y (in % of GDP in 2003) and political orientation x (median value for country, measured as an index from 1 to 10, with 10 indicating the right-wing orientation)

$\hat{\alpha} = 41.099$, $\hat{\beta} = 3.444$, $\bar{R}^2 = 0.0506$, t statistic is -1.063 (not significant). Data Source: OECD, World Values Survey.

provide supporting evidence, which however is weak to begin with in view of the very low coefficient of determination.

As a sixth and final behavioral dimension, one can cite religiosity. There are three strands of theory, all of them predicting a negative relationship between religiosity and income redistribution. First, Benabou and Tirole (2006) model collective cultural beliefs, one of which is religion. In their 'highly religious' (Protestant) equilibrium, hard work and industriousness are believed to have rewards in the afterlife, the amount of redistribution is low, and average effort and output are high. In their 'less religious' equilibrium, there is less effort and more redistribution (e.g. through alms). Second, Scheve and Stasavage (2006a,b) argue and pro-

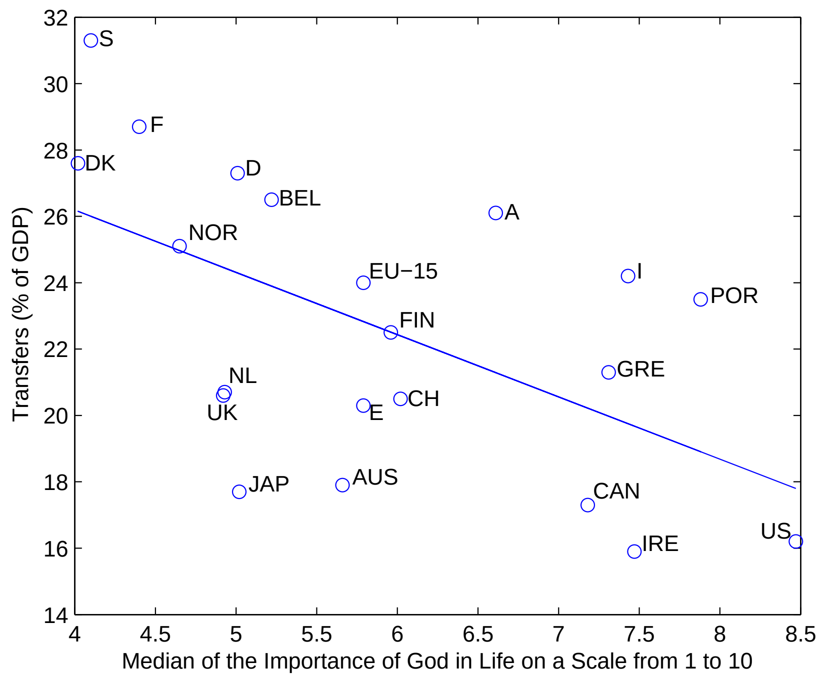


Figure 2.10: Relationship $y = \hat{\alpha} + \hat{\beta}x$ between transfers y (in % of GDP in 2003) and religiosity x (median value for country, measured as an index from 1 to 10, with 10 indicating the highest importance of God in life)

$\hat{\alpha} = 33.695$, $\hat{\beta} = -1.877$, $\bar{R}^2 = 0.3175$, t statistic is -2.906 (significant at the 1 percent level). Data Source: OECD, World Values Survey.

vide evidence that religion provides insurance against adverse events. Therefore risk-averse religious individuals express less demand for redistribution as a collective insurance device, resulting again in a negative predicted relationship between religiosity and redistribution. A third strand argues that public welfare crowds out participation in church and charitable activities, giving once more rise to a negative correlation. Hungerman (2005) and Gruber and Hungerman (2007) find evidence that public insurance spending indeed crowds out religious charitable spending. Figure 2.10 shows the strength of religious orientation (1 = no importance of God in life, 10 = maximum importance) to vary considerably, with the United States marking the high end. With a coefficient of determination of 0.32 and a t statistic of -2.906, this is the best-fitting bivariate regression designed to explain the share of transfers in the GDP. Hence, the partial correlation between religiosity and the share of transfers in GDP is clearly negative, supporting the theories expounded above. On this score, Switzerland shares the somewhat guarded attitudes prevailing in the EU. Being located close to the regression line, it provides additional evidence supporting the theoretical arguments relating religion to redistribution.

2.6 Final Assessment

In sum, out of six behavioral factors that according to the existing literature influence attitudes with regard to income redistribution, all (with the exception of ethno-linguistic fragmentation) were found to be at least partially correlated with the amount of income distribution as measured by public transfers as a share of GDP. And in all cases, Switzerland, located between the United States and the EU, is on or close to the regression line, in contradistinction to the economic and political explanations considered. This observation is informative: Switzerland consistently lies between the United States and the EU average on all six scales used as explanatory variables. If the estimated relationships have validity, it should therefore be located on or close to the regression line rather than constituting an outlier. Since this prediction is confirmed, it tips the balance in favor of behavioral explanations of income redistribution.

A final assessment can be based on multivariate analysis relating the share of transfers in GDP to mobility, openness (both economic), proportional representation (political), ethno-linguistic fragmentation, migration rate, birth rate, belief that luck determines success, political orientation, and religiosity (all behavioral). Applying the stepwise reduction procedure by excluding the least significant regressors, one arrives at the final model of Table 2.8. The two explanatory variables retained are both behavioral, viz. the belief that luck determines success and religiosity. Moreover, their coefficients do not significantly differ from the coefficients in the respective bivariate regressions. They are both significant at the 5 percent significance level in the final model as shown in Table 2.8 (compared to a 1 percent significance level in the bivariate regressions, cf. Figures 2.8 and 2.10). However, these two variables are jointly significant at the 1 percent level, as indicated by the test statistic $F(2, 17) = 7.19$ [0.0055].

	coefficient	standard error	t value
constant	19.290	7.925	2.434
belief in luck	2.537	1.215	2.089
religiosity	-1.457	0.654	-2.229

Joint significance test: $F(2, 17) = 7.19$ [0.0055]
Table 2.8: Final model for the share of public transfers in GDP

2.7 Conclusion

In this paper, an attempt was made to explain the difference in the amount of public income redistribution between the United States and the European Union (EU), based on an empirical examination of three sets of determinants, economic, political, and behavioral, with the most recent data for 20 developed OECD countries[6] listed in Figure 2.1. The previous literature [Alesina and Glaeser (2004)] looks at bivariate relations between the amount of public redistribution and various economic, political, and behavioral variables for large data sets including developing countries. In addition to the variables in Alesina and Glaeser (2004), we included further variables in our study such as a measure for social mobility, migration rate and birth rate. Since Switzerland, a non-EU country, is almost always located between the two polar cases, we use it as a test case providing corroborating or contradicting evidence.

Economic determinants predict more rather then less income redistribution in the United States than in EU, contrary to facts. Before-tax income inequality is higher, the income distribution is more skewed, and incomes and terms of trade are more volatile in the United States than in the EU countries. However, U.S. income mobility is higher, too, possibly serving as a substitute for redistribution. Pertinent bivariate regressions have poor statistical fit. Moreover, Switzerland lies rather far from the respective regression lines, providing contradicting rather than corroborating evidence.

Political variables include district rather than proportional representation, a two-party vs. multiparty system, a presidential vs. parliamentary democracy, courts emphasizing property rights, and failure of a strong and lasting socialist party to form; all distinguishing the United States from the EU. However, once again, the bivariate regressions do not have much explanatory power. And again, Switzerland comes close to being an outlier, thus failing to buttress the weak supporting evidence.

[6] However, the regression on the belief about luck vs. effort as well as the final multivariate regression do not include Greece due to a lack of data.

Behavioral explanations include ethno-linguistic fragmentation of the country, the migration rate, the birth rate, the belief that luck determines success, the degree of left-wing orientation, and the strength of religious belief. On several of these scores, the U.S. population constitutes an outlier. In particular, it sees hard work rather than luck as a determinant of success, contrary to the population of a typical EU country. Two bivariate regressions (with belief that the luck determines economic success and religion as the explanatory variable, respectively) attain coefficients of determination of 0.3 or more. In addition, the Swiss observation is on or close to the regression line, thus providing supporting evidence. In a final assessment, we identify the most significant variables based on a multivariate regression, complementing the bivariate analyses by Alesina and Glaeser (2004).

Both the bivariate and the multivariate regressions suggest the following conclusions. The United States has less income redistribution than the European Union for three main reasons. The first is political. With its absence of proportional representation (a feature shared with Australia, Canada, and the United Kingdom, countries with a low amount of redistribution, too), the United States has an impediment against resolving political conflict through buying off minorities, a tradition characterizing notably Austria and Sweden (see Figure 2.4 again). Using Switzerland as a test case, its observation is off the regression line by about the same amount as the United States (and on the same side). Therefore, it does contribute a measure of confirmatory evidence. The other two reasons are behavioral. The U.S. population does not believe that chance determines economic success, contrary to the EU population (see Figure 2.8 again). Further, it believes that God is of critical importance in life, which is held to a comparable degree by the Portuguese but certainly not by the EU population on average (see Figure 2.10 again). On both scores, the Swiss observation is on or close to the respective regression line, providing a bit of supporting evidence. And on both scores, Switzerland is located between the United States and the EU, showing its Janus face.

It is appropriate to point out the limitations of this analysis. First, it does not rest on a unifying theoretical basis, drawing on economics, political science, and sociology in an eclectic manner. Second, possible determinants are tested mainly one by one in a series of bivariate regressions. This of course entails the risk of attributing influence to a factor that should be attributed to another factor not controlled for. Third, the evidence relates to a point in time. Measured values can be subject to transitory shocks causing them to differ from the permanent values the theories refer to. Fourth, one could argue that while accepting the view that the United States and the European Union constitute two polar cases with regard to income redistribution, some country other than Switzerland should have been selected as a test case in between. All these limitations have to be taken seriously. Above all, they call for additional research to answer the question, "Why is there such a marked difference between the United States and the European Union in terms of income redistribution?" The present study may provide a few preliminary answers that need to be corroborated. It uses Switzerland as a test case because that country, while being in the middle of Europe, does have a few features that are reminiscent of the United States, giving it an intriguing Janus face.

2.8 Acknowledgements

The authors gratefully acknowledge financial support from the Swiss National Science Foundation (SNF) under Project no. 100012-116398. They received helpful comments from Rosa Fontes, Philipp Morf, Maurus Rischatsch, Mercedes Sastre, Michèle Sennhauser, and Philippe Widmer, participants in the 16th Symposium on Public Economics (5-6 February 2009, Granada, Spain), and three anonymous referees, and are grateful to Georgios Sismanidis for help with data collection.

Economic well-being, social mobility, and preferences for income redistribution: evidence from a discrete choice experiment

Ilja Neustadt and Peter Zweifel

Submitted to “International Tax and Public Finance”.

Abstract: In this paper, preferences for income redistribution in Switzerland are elicited through a Discrete Choice Experiment (DCE) performed in 2008. In addition to the amount of redistribution as a share of GDP, attributes also included its uses (working poor, the unemployed, old-age pensioners, families with children, people with ill health) and nationality of beneficiary (Swiss, Western European, others). Willingness to pay for redistribution increases with income and education, contradicting the conventional Meltzer-Richard (1981) model. The Prospect of Upward Mobility hypothesis [Hirschman and Rothschild (1973); Benabou and Ok (2001)] receives partial empirical support.
Keywords: Income redistribution, preferences, willingness to pay, discrete choice experiments, stated choice, economic well-being, social mobility

JEL classification: C35, C93, D63, H29

Chapter 3

Economic well-being, social mobility, and preferences for income redistribution: evidence from a discrete choice experiment

3.1 Introduction

Politicians and interest groups often claim to know citizens' preferences with regard to income redistribution. While the typical right-wing stance is to decry it as excessive, the left points to pockets of poverty even in rich societies that need to be eradicated through more redistribution. The economists' contribution to the debate traditionally has been to analyze the effects of redistributive policies on employment, output, and growth. This paper intends to go a step further by measuring citizens' willingness to pay (WTP) for redistribution. Through a Discrete Choice experiment (DCE), it seeks to determine not only the desired amount of redistribution but also to test several hypotheses concerning the determinants of this WTP. The data come from a DCE performed in the fall of 2008 and involving 979 Swiss citizens.

Recently, there has been a great deal of research into the demand for redistribution and its determinants, which will be discussed in detail in Section 3.2 below. One strand relates the measured amount of redistribution to economic, institutional, and behavioral factors.

Examples are Alesina and Giuliano (2009) and Akkoyunlu et al. (2009). However, the observed amount of redistribution is the outcome of an interaction between demand and supply, with supply governed by a country's political institutions and processes. This classical identification problem would have to be addressed in order to make inference about citizens' preferences for redistribution. A second strand of research, exemplified by Alesina and La Ferrara (2005) and Guillaud (2008), relies on surveys designed to measure attitudes towards redistribution. The problem with this approach is its failure to impose a budget constraint. It therefore cannot predict actual decision making (e.g. voting at the polls), where citizens take into account the consequences in terms of their own income and wealth. A third approach seeks to solve this problem through Contingent Valuation (CV) experiments [see e.g. Boeri et al. (2002)]. The weakness of the CV approach is that it holds all the attributes of the good in question constant, varying its price only. In the present context, one would want to vary other attributes of redistribution besides its tax price, viz. its use (for health, old age, etc.) and the type of beneficiary (foreigner, national).

By way of contrast, a DCE allows to measure preferences uncontaminated by supply influences, it imposes the budget constraint through the price attribute, and it does so in a realistic way by making respondents choose between alternatives where all attributes are allowed to vary.

There are two recent contributions whose methodology is similar to the one adopted in this paper. One is by Andreoni and Miller (2002), who test the consistency of altruistic revealed preferences in a dictatorship experiment, varying an implicit price. Their method of inferring preferences through estimating WTP values is close to this paper. The other is by Kuhn (2005), who asked Swiss respondents to estimate wages earned by different professions as well as indicate the wages they deemed fair. The difference between these two values was then used as an indicator of the demand for redistribution. On average, preferences were for the wages of high-earning professions such as lawyers, physicians or federal ministers to be reduced by 10 percent while those of low-income groups, to be increased by some 5 percent. Interestingly, such a redistributive scheme would roughly result in budget balance.

The remainder of this paper is structured as follows. Section 3.2 contains a literature review from which hypotheses to be tested are derived. Its first part concerns the general determinants of the demand for redistribution, the second, economic well-being, and the third, mobility as determinants of preferences for redistribution. Section 3.3 presents a general description of the method of DCEs as well as the design of the present experiment. The descriptive statistics of the experiment follow in Section 3.4, and hypothesis tests, in Section 3.5. Section 3.6 summarizes the results and concludes with implications for public policy.

3.2 Literature review and statement of hypotheses

This section first presents research that defines the general background of this paper and then moves on to contributions that lead to a set of specific hypotheses to be tested.

3.2.1 General determinants of the demand for income redistribution

In their reviews, Alesina and Giuliano (2009) and Akkoyunlu et al. (2009) [see Chapter 2 of this dissertation] identify a wide set of factors influencing preferences that can be categorized as economic, political, and behavioral determinants. As to the economic determinants, Alesina and La Ferrara (2005) empirically analyze the effects of current and future income on the demand for redistribution in the United States. While low current income bolsters demand, chances for higher future income reduce it when the tax system is expected to become more progressive. Another economic explanation, suggested by the social contract literature, is that a preference for redistribution can at least in part be interpreted as demand for insurance by risk-averse individuals. In a hypothetical situation, where individuals do not yet know their endowment as well as their future position in society [‘veil of ignorance’, cf. Rawls (1999)], they exhibit a positive WTP for an income transfer from more favorable future states to less favorable ones. Redistributive policies can thus be interpreted as reflecting this hypothetical demand for insurance.

Beck (1994) investigates individual behavior under the ‘veil of ignorance’ in an experiment. Placing participants in a hypothetical society with random differences in income, represented by lotteries, he derives the desired amount of income redistribution. Individuals indeed display risk aversion, albeit not of the extreme kind implied by the Rawlsian maximin rule[1]. Furthermore, they show no preference for income redistribution in excess of what can be explained by risk aversion.

[1]The Rawlsian maximin rule uses the maximum improvement of the individual with minimum initial wealth as the sole criterion.

As to the political determinants, the literature [Persson and Tabellini (2000, 2003), Lizzeri and Persico (2001), Milesi-Ferretti et al. (2002)] predicts that proportional representation tends towards universal programs benefitting various groups (old-age pensioners, working poor, minorities, etc.), while majority rule results in targeted 'pork barrel' programs. Persson and Tabellini (2003) find supporting empirical evidence in that countries with proportional representation have GDP shares of government expenditure that ceteris paribus are 5 percentage points higher than with majority rule. Moreover, Akkoyunlu et al. (2009) [Chapter 2] show that there is a weak evidence of a positive correlation between the degree of proportional representation and the transfer share in GDP in OECD countries. Additional political determinants of redistribution include two-party vs. multiparty system, presidential vs. parliamentary democracy, and direct vs. representative democracy, with two-party systems, presidential, and direct democracies all predicted to induce less public redistribution. Switzerland on the one hand has a high degree of proportional representation and a parliamentary democracy; on the other hand, its extensive direct democratic control might serve to limit public welfare spending while enforcing efficiency in redistribution [cf. Feld et al. (2007)].

Among the behavioral determinants of income redistribution, beliefs have been at the center of attention. The theoretical base is laid by Alesina and Angeletos (2005), who develop a model where society's belief whether effort or luck determines economic success gives rise to multiple self-fulfilling equilibria; Benabou and Tirole (2006) propose a model for the emergence and persistence of such collective beliefs. On the empirical side, Fong (2001) presents evidence in line with Alesina and La Ferrara (2005) suggesting that beliefs about the role of luck in determining economic success are an important determinant of the demand for redistribution. She also considers the effects of incentives. If effort determines income, then an increased income tax rate causes a loss in output due to its effect on incentives. This consideration is hypothesized to qualify the link between beliefs and the demand for redistribution. However, the data fail to support this hypothesis.

Boeri et al. (2001) study international attitudes towards redistribution with a focus on pension and unemployment schemes in France, Germany, Italy, and Spain. They also perform CV experiments that impose an explicit trade-off between income and social insurance coverage on respondents. They find that people oppose an extension of the welfare state, with conflicts between young and old, rich and poor, and insiders and outsiders creating significant hurdles to welfare reform.

3.2.2 Economic well-being and demand for income redistribution

The standard model of income redistribution, originally proposed by Romer (1975) and Roberts (1977) and extended by Meltzer and Richard (1981) [RRMR model], assumes that identical non-altruistic utility-maximizing individuals are only differentiated by their income levels and determine their individually optimal consumption and leisure. The utility function of individual i takes the following quasi-linear form [cf. Persson and Tabellini (2000)],

$$u_i(c_i, l_i) = c_i + v(l_i) \tag{3.1}$$

where c_i denotes individual consumption, l_i leisure, and $v(\cdot)$ is an increasing and concave function. The government pays a lump-sum transfer T to all citizens, which is financed by a linear uniform income tax τ. Thus, the household budget constraint takes the form

$$c_i + (1-\tau)l_i \quad \leq \quad (1-\tau)(\omega + y_i) + T \tag{3.2}$$

with ω denoting the household's time endowment and y_i individual productivity[2], distributed in the population according to a distribution function $F(\cdot)$ with $E[y_i] = \mu$ and $\text{Med}[y_i] = m < \mu$. Solving the utility maximization problem yields the optimal demand function for leisure given by

$$\hat{l}_i \quad = \quad v_l^{-1}[1-\tau], \tag{3.3}$$

with v_l denoting i's marginal utility of leisure (subscript i dropped for simplicity). The government's budget constraint reads

$$T \leq \tau \int_{y_i} (\omega + y_i - l_i) dF(y_i). \tag{3.4}$$

[2] y_i can be alternatively interpreted as (i) personal income before tax, (ii) personal level of education, or (iii) subjective self-positioning on a social distance scale.

The utility-maximizing tax rate $\hat{\tau}_i$ for individual i is thus implicitly given by

$$\hat{\tau}_i \quad = \quad (y_i - \mu)\, v_{ll}[\hat{l}_i[\hat{\tau}_i]]. \tag{3.5}$$

By concavity of $v(\cdot)$ ($v_{ll} < 0$), individuals with an income below the mean favor taxation and transfers while individuals with an income above the mean oppose it. In a political equilibrium, the majority of voters supports a positive tax rate that corresponds to the value $\hat{\tau}_m = (m - \mu)\, v_{ll}[\hat{l}_i[\hat{\tau}_m]]$ desired by the median voter, whose income is assumed to be below the mean (which holds for most economies). The model's prediction is that the more unequal the income distribution, i.e. the larger the gap between the mean and the median income, the higher the level of taxation and redistribution.

The empirical evidence is quite mixed. On the one hand, Alesina and Rodrik (1994), Persson and Tabellini (1994), and Milanovic (2000) find some supporting evidence. Furthermore, Guillaud (2008), conducting a cross-section analysis of survey data from four EU countries, shows that poorer and less educated individuals are more in favor of redistribution. On the other hand, Alesina and Glaeser (2004), Perotti (1996), and Rodriguez (1999) fail to find supporting evidence for this model.

Based on the RRMR model, we can formulate the static Hypothesis 1 relating the demand for income redistribution to alternative measures of the individual's current economic well-being, viz. personal income, personal level of education, or self-positioning on a social distance scale, respectively.

Hypothesis 1: *The demand for redistribution is expected to decrease with*

(a) *personal income,*

(b) *personal level of education,*

(c) *higher self-positioning on a social distance scale.*

3.2.3 Social mobility and demand for income redistribution

The idea that attitudes toward public redistribution could be explained by individuals' mobility was originally introduced by de Tocqueville (1835). More recently, Piketty (1995) considered a model of learning from income mobility experience and explained persisting differences in attitudes towards redistribution. In the long run, those who experienced upward mobility tend to believe more in effort and demand less redistribution.

This "Prospect of Upward Mobility" (POUM) hypothesis, originally suggested by Hirschman and Rothschild (1973) as the 'tunnel effect' and more recently reformulated by Benabou and Ok (2001), extends the RRMR model by introducing individuals' expectations, based on their observations regarding the income mobility of others in society. Thus, upward mobility may dampen a poor but forward-looking voter's enthusiasm for income redistribution. The three premises for this result are: (i) future expected income is a concave function of current income, (ii) individuals are not too risk averse, and (iii) the government is committed to an unchanged fiscal policy.

In a simplified version, the Benabou-Ok model can be illustrated by the following two-period example. Suppose that tomorrow's income y_1 is a concave function of today's income y_0: $y_1 = f(y_0)$ with $f''(y) < 0$ for all $y \in [0, y^{\max}]$. Function $f(\cdot)$ is normalized such that the individual with the mean income μ_0 today earns exactly the same income tomorrow, $\mu_0 = f[\mu_0]$. Then agents with current income below average expect a higher income tomorrow while those above average will expect a decline of income. By concavity of $f(\cdot)$, total income gains of the poor are smaller than total losses of the rich. Thus, tomorrow's average income μ_1 must fall short of today's average μ_0. Therefore, all individuals with current incomes in the interval $(f^{-1}(\mu_1), \mu_0)$ expect their future income to be higher than average μ_1 and thus oppose redistribution in the next period.

Empirical support of the POUM hypothesis is provided by Alesina and La Ferrara (2005) who, using an actual mobility matrix for the United States, show that people who expect high future income oppose redistribution. This 'tunnel effect' also works in the opposite direction, causing forward-looking agents with high incomes but downward mobility expectations to be in favor of redistribution. This prediction is confirmed by Ravallion and Lokshin (2000) using a data set from Russia. Furthermore, Molnár and Kapitány (2006a,b) show that individuals who lack clear expectations about their future income favor redistribution even more so than those with negative but clear expectations. Rainer and Siedler (2008) use probabilistic expectations data to show that individuals with a sufficiently large chance of occupational upward mobility exhibit a lower demand for redistribution; conversely, those with a sufficiently large risk of occupational downward mobility opt for more redistribution. Checchi and Filippin (2004), testing the POUM hypothesis by means of a within-subjects experiment, find corroborating evidence under several alternative specifications.

According to Guillaud (2008), however, individuals who subjectively experienced upward mobility over ten years tend to be more (rather than less) supportive of redistributive policies. Moreover, upward intergenerational mobility (measured as the difference in the job prestige compared to the job of the father) leads to a more positive rather than negative attitude towards redistribution. Alesina and Giuliano (2009) review the theoretical literature, providing a framework for incorporating various effects that were previously studied in isolation. They examine the empirical evidence for the United States and briefly across countries, concluding that social mobility (if measured as the change in the occupational prestige) does decrease demand for redistribution once sociodemographic (age, gender, race) and socioeconomic characteristics (income, education) are controlled for.

Based on the POUM hypothesis, we formulate the dynamic Hypothesis 2 relating the demand for redistribution to alternative mobility measures, viz. difference in education between individuals and their fathers, difference in the occupational prestige between individuals and their fathers (both intergenerational mobility), past income mobility, expected income mobility, as well as the experienced change in the self-positioning on a social distance scale (subjective mobility).

HYPOTHESIS 2: *The demand for redistribution is expected to decrease with*

(a) *a higher difference between individuals and their fathers in terms of education,*

(b) *a higher difference between individuals and their fathers in terms of occupational prestige,*

(c) *higher upward income mobility in the past,*

(d) *higher upward income mobility in the future,*

(e) *larger positive change in the self-positioning on a social distance scale.*

3.3 Discrete choice experiments

3.3.1 Theoretical foundations

Discrete Choice Experiments (DCEs) provide a tool for measuring individuals' preferences for characteristics of commodities, the so-called attributes. In contradistinction with classical Revealed Preference Theory, originating with Samuelson (1938), DCEs allow individuals to express their preferences for non-marketed as well as hypothetical products. During a DCE, respondents are repeatedly asked to compare the status quo with several hypothetical alternatives defined by their attributes including their price. By varying the levels of attributes, different product alternatives are generated. A rational individual will always choose the alternative with the highest utility level. From the observed choices, the researcher can infer the utility associated with the attributes. The proposed method, derived from the New Demand Theory of Lancaster (1971), is also known as Conjoint Analysis [Louviere et al. (2000)].

The most prominent alternative to a DCE is Contingent Valuation (CV). A certain situation or product is described in detail and respondents are asked to indicate their maximum willingness to pay (WTP) for this fixed product. Only its price attribute is varied, while in Conjoint Analysis all relevant attributes are varied simultaneously, making it a multi-attribute valuation method [Merino-Castello (2003)]. While a DCE describes the product in less detail than a typical CV study, it allows for analyzing many product varieties by varying the levels of relevant attributes [cf. Louviere et al. (2000), p. 344]. Trade-offs among attributes can be explicitly taken into account and WTP values of attributes estimated separately (see below). Furthermore, strategic behavior of respondents is less likely than in CV with its exclusive emphasis on price, which facilitates strategic behavior. Finally, biases that easily occur when individuals are directly asked about their WTP are less frequently observed in a DCE [Ryan (2004)].

A particular advantage of a DCE in the present context is that it permits to explicitly impose the budget constraint through a price attribute in the guise of the tax share of income used to finance the transfers considered. Respondents can be made to simultaneously choose this share and hence the 'size of the pie' and the 'slices of the pie' devoted to different types of recipients and uses (health, old age, etc.). Thus, trade-offs among different attributes of the redistribution plan can be calculated to assess the relative importance of the respective redistributive goals.

The econometric method used is based on the Random Utility Theory [see Luce (1959), Manski and Lerman (1977), McFadden (2001), McFadden (1974, 1981, 2001)]. Individual i values alternative j according to the utility V_{ij} attained, which is given by

$$V_{ij} = v_i(a_j, p_j, y_i, s_i, \varepsilon_{ij}). \tag{3.6}$$

Here, $v_i(\cdot)$ denotes i's indirect utility function, a_j, the amount of attributes associated with alternative j, and p_j, price. The individual's income and sociodemographic characteristics are symbolized by y_i and s_i, respectively. Finally, ε_{ij} denotes the error term, which is due to the fact that the experimenter will never observe all the arguments entering v_i, imparting a stochastic element to observed choices.

As usual, the utility function is additively split into a systematic component $w(\cdot)$ and a stochastic one,

$$V_{ij} = w_i(a_j, p_j, y_i, s_i) + \varepsilon_{ij}. \tag{3.7}$$

A utility-maximizing individual i will prefer alternative j to alternative l if and only if

$$w_i(a_l, p_l, y_i, s_i) + \varepsilon_{il} \leq w_i(a_j, p_j, y_i, s_i) + \varepsilon_{ij}. \tag{3.8}$$

Due to the presence of the stochastic term, only the probability P_{ij} of individual i choosing alternative j rather than alternative l can be estimated, with

$$P_{ij} = \text{Prob}\left[w_i(a_l, p_l, y_i, s_i) + \varepsilon_{il} \leq w_i(a_j, p_j, y_i, s_i) + \varepsilon_{ij}\right] \tag{3.9}$$

$$= \text{Prob}\left[\varepsilon_{il} - \varepsilon_{ij} \leq w_i(a_j, p_j, y_i, s_i) - w_i(a_l, p_l, y_i, s_i)\right]. \tag{3.10}$$

Thus, the probability of choosing j amounts to the probability of the systematic utility difference $w_i[j] - w_i[l]$ dominating the 'noise', $\varepsilon_{il} - \varepsilon_{ij}$. The error terms $\{\varepsilon_{il},\ \varepsilon_{ij}\}$ can be assumed to be normally distributed with mean zero and variances σ_l^2 and σ_j^2 as well as covariance σ_{lj}. Under these assumptions, $\varphi_{ij} := \varepsilon_{il} - \varepsilon_{ij}$ is also normally distributed with mean zero and variance $\sigma^2 := \mathrm{Var}[\varphi_{ij}] = \sigma_l^2 + \sigma_j^2 - 2\sigma_{lj}$. Thus, equation (3.10) can be represented as

$$P_{ij} = \Phi\left(\frac{w_i(a_j, p_j, y_i, s_i) - w_i(a_l, p_l, y_i, s_i)}{\sigma}\right), \tag{3.11}$$

where $\Phi(\cdot)$ denotes the cdf of a standard normal distribution. The model is known as the binary probit model [cf. Ben-Akiva and Lerman (1985)]. Hensher et al. (1999) provide empirical evidence that a linear specification of the function $w(\cdot)$ leads to good predictions in its middle ranges. Therefore, one posits

$$w_i(a_j, p_j, y_i, s_i) = c_i + \sum_{k=1}^{K} \beta_k a_k + \varepsilon_{ij}, \tag{3.12}$$

where c_i represents an individual-specific constant, a_k, $k = 1, \ldots, K$, are the attributes of the alternative, and β_k, $k = 1, \ldots, K$, are the parameters to be estimated. These parameters can be interpreted as the constant marginal utilities of the corresponding attributes. The marginal rate of substitution between two attributes m and n is given by

$$\mathrm{MRS}_{m,n} = -\frac{\partial v/\partial a_m}{\partial v/\partial a_n}. \tag{3.13}$$

In the case of a linear utility function, this can be estimated as the ratio of the respective slope parameters,

$$\mathrm{MRS}_{m,n} = -\frac{\hat{\beta}_m}{\hat{\beta}_n},$$

representing the marginal WTP for an additional unit of a_m expressed in units of a_n. Therefore, the marginal WTP for attribute a_m can be calculated by dividing the marginal utility of this attribute by the marginal utility of the price attribute [in the present context, the income tax rate, see e.g. Telser (2002), p. 56]:

$$\mathrm{MWTP}(a_m) = \frac{\partial v/\partial a_m}{\partial v/\partial p_j}. \tag{3.14}$$

Notice that, by Roy's Identity, $x_{ij} = -\frac{\partial v(\cdot)/\partial p_j}{\partial v(\cdot)/\partial y_i}$, the (uncompensated) demand of individual i for commodity j corresponds to the negative ratio of partial derivatives of the indirect utility function with respect to price p_j and income y_i. If one alternative is chosen, then the optimal quantity demanded is equal to one, i.e. $x_{ij} = 1$. Therefore, Roy's Identity yields $\frac{\partial v}{\partial y_i} = -\frac{\partial v}{\partial p_j}$, i.e. the marginal utility of income is equal to the negative derivative of the indirect utility function with respect to price.

By limiting the specification to the product attributes only (simple model, cf. Section 3.5.1), one obtains the following expression representing the difference in utility of individual i between alternative j and the status quo,

$$\Delta V_{ij} \quad = \quad c_i + \sum_{k=1}^{K} \beta_k \Delta a_{kj} + \beta_p \Delta p_j + \varphi_{ij}, \tag{3.15}$$

where $\Delta a_{kj} = a_{kj} - a_{lj}$, $\Delta p_j = p_j - p_l$, $c_i = c_{il} - c_{ij}$, and $\varphi_{ij} = \varepsilon_{il} - \varepsilon_{ij}$ for each $j \neq l$. This simple model suffices to estimate WTP values of an average respondent (see Section 3.5.1).

For econometric inference, it is important to take into account that the same individual makes several choices. The two-way random-effect specification takes this into account with $\varphi_{ij} = \mu_i + \eta_{ij}$, where μ_i denotes the component that varies only across individuals but not across the choice alternatives. The terms μ_i and η_{ij} are assumed uncorrelated with the product attributes $(a_{i1}, \ldots, a_{iK})$ and between themselves. By a standard assumption in a probit model, $\sigma_\eta = 1$. Hence $\mathrm{Var}[\varphi_{ij}] = \sigma_\eta^2 + \sigma_\mu^2 = 1 + \sigma_\mu^2$ and $\mathrm{Corr}[\varphi_{ij}, \varphi_{il}] = \frac{\sigma_\mu^2}{1+\sigma_\mu^2} =: \rho$. The parameter ρ indicates how strongly the various responses of an individual are correlated with each other, or, equivalently, the share of the total variance that can be explained by the individual-specific error term. The random-effects specification is justified if ρ is high and significant.

The simple model can be extended by including various socioeconomic variables (e.g. income group, level of education, social mobility). These variables need to be interacted with the product attributes as well as with the constant, giving rise to the extended model specification which allows to check for preference heterogeneity and thus to test Hypotheses 1 and 2, cf. Section 3.5.2. By means of a t test we can investigate whether the differences in marginal WTP values between different socioeconomic groups are statistically significant. The computation of the variance of the marginal WTP values can be performed by the delta method, cf. Hole (2007). The estimate of the variance is given by

$$\begin{aligned}
\mathrm{Var}\left[-\frac{\hat{\beta}_k}{\hat{\beta}_p}\right] &= \left[\frac{\partial\left(-\hat{\beta}_k/\hat{\beta}_p\right)}{\partial\hat{\beta}_k}\right]^2 \mathrm{Var}[\hat{\beta}_k] + \left[\frac{\partial\left(-\hat{\beta}_k/\hat{\beta}_p\right)}{\partial\hat{\beta}_p}\right]^2 \mathrm{Var}[\hat{\beta}_p] \\
&\quad - 2\,\frac{\partial\left(-\hat{\beta}_k/\hat{\beta}_p\right)}{\partial\hat{\beta}_k}\frac{\partial\left(-\hat{\beta}_k/\hat{\beta}_p\right)}{\partial\hat{\beta}_p}\mathrm{Cov}[\hat{\beta}_k,\hat{\beta}_p] \\
&= \frac{1}{\hat{\beta}_p^2}\,\mathrm{Var}[\hat{\beta}_k] + \frac{\hat{\beta}_k^2}{\hat{\beta}_p^4}\,\mathrm{Var}[\hat{\beta}_p] + 2\,\frac{\hat{\beta}_k}{\hat{\beta}_p^3}\,\mathrm{Cov}[\hat{\beta}_k,\hat{\beta}_p].
\end{aligned}$$

3.3.2 Experimental design

In order to elicit the preferences of Swiss citizens for income redistribution, a representative telephone survey with 979 respondents was conducted in the fall of 2008. Prior to the telephone survey, the attributes and their levels used to define 'income redistribution' had been checked in two pretests for their relevance. They form four groups (see Table 3.1).

1. Shares of the total redistribution budget (to be spent on five types of recipients, viz. the working poor, the unemployed, old-age pensioners, families with children, and people with ill health);
2. Shares of the total redistribution budget (to be spent on three groups, viz. Swiss citizens, western European foreigners, and other foreigners);
3. Total amount of redistribution, defined as a share of GDP;
4. Share of personal income tax rate to be paid by the respondent (the price attribute).

Clearly, these attributes and their levels combine to form a total number of possible scenarios that cannot be realized in an experiment. The scenarios define the n rows of the observation matrix X, with associated covariance matrix $\Omega = \sigma^2\left(X'X\right)^{-1}$ of parameters β to be estimated.

So-called D-efficient design calls for the minimization of the geometric mean of the eigenvalues of Ω,

$$D\text{ efficiency} = \left(|\Omega|^{\frac{1}{K}}\right)^{-1}$$

Attribute	Label	Status Quo Level	Alternative Levels
Shares of benefits going to			
• Working Poor	W_POOR	10%	5%, 15%
• Unemployed	UNEMP	15%	5%, 25%
• Old-Age Pensioners	PENS	45%	35%, 55%
• Families with Children	FAM	5%	10%
• Ill People	ILL	25%	20%, 30%
Shares of benefits going to			
• Swiss citizens	SWISS	75%	60%, 85%
• Western European foreigners	WEU_FOR	10%	5%, 20%
• Other foreigners	OTH_FOR	15%	10%, 20%
Total amount of redistribution	REDIST	25% (of GDP)	10%, 20%, 30%, 40%, 50%
Income tax	TAX	25% (of personal income)	10%, 15%, 40%

Table 3.1: Attributes and their levels

where K denotes the number of parameters to estimate [cf. Carlsson and Martinsson (2003)]. Using this optimization procedure and incorporating several restrictions, the number of alternatives was reduced to 35 and randomly split in five groups. One alternative was included twice in each decision set for a consistency test, resulting in 8 binary choices per respondent. In order to make sure that decisions were based on a homogeneous information set and made in a consistent way, respondents were provided with a detailed description of the attributes and their possible realizations. The Appendix shows the graphical representation of the status quo (Exhibit 1) and two selected alternatives (Exhibits 2 and 3).

3.4 Descriptive statistics

3.4.1 Socioeconomic characteristics

The sample consists of 979 respondents, 70 percent of them residing in the German-speaking part and 30 percent in the French-speaking part of Switzerland. Some 94 percent are born in the country, 50 percent are men, 20 percent having a monthly income below CHF 2,000 and 23 percent, above CHF 6,000, reflecting the structure of the Swiss population. However, only 1.5 percent of the respondents are unemployed.

42.6 percent of the respondents agreed with the statement, 'By increasing the income tax rates for rich families and financially supporting poor families, the government should try to reduce the income gap between rich and poor.' while 54.6 percent disagreed. On the other hand, 36 percent of the respondents stated that the current level of social benefits was too low, 9 percent stated that it was too high, and 48.7 percent found it exactly right.

	Current		5 years ago		In 5 years	
Income classes, CHF	No.	% of valid answers	No.	% of valid answers	No.	% of valid answers
< CHF 2000	192	20	236	25	135	14
CHF 2000 - 3999	193	20	189	20	187	20
CHF 4000 - 5999	344	36	300	32	349	37
≥CHF 6000	221	23	223	23	264	28
Total valid answers	950	100	948	100	935	100
Missing	29		31		44	
Sample	979		979		979	

Table 3.2: Current, past, and future expected individual incomes, per month (in CHF)

The frequency distributions of current, past, and expected future incomes are shown in Table 3.2. Note that incomes <CHF 2000, CHF 2000-3999, and ≥CHF 6000 approximately correspond to the first, second, and fifth income quintiles whereas the bracket CHF 4000-5999 contains the third and the fourth quintiles. From the individual responses entered in Table 3.2, transition probabilities between the income quintiles were estimated (which are not available from official Swiss statistics).

Table 3.3 shows the frequency distributions of the respondents' own as well as their fathers' educational levels.

Table 3.4 contains the frequency distribution of the differences between the respondents' and fathers' educational levels, which will be referred to as DIFF_ED, as well as the distribution of answers to the question, 'Is there a difference in occupational prestige in the society between your job and your father's job?', later referred to as DIFF_PREST. This is an indicator of subjective intergenerational mobility INTERG_MOB_SUBJ.

Table 3.5 shows the current and future expected self-positioning of respondents on a social distance scale. Using these two variables, one can determine the distribution of the subjectively expected social mobility to occur within a generation.

	Respondents		Fathers	
Educational level	No.	% of valid answers	No.	% of valid answers
Less than high school	654	67	670	69
High school	195	20	185	19
College and more	129	13	111	11
Total valid answers	978	100	966	100
Missing	1		13	
Sample	979		979	

Table 3.3: Respondents' and fathers' educational levels

	Education		Occupational prestige	
Difference	No.	% of valid answers	No.	% of valid answers
Positive	194	20	331	35
No difference	600	62	361	38
Negative	172	18	138	15
Total valid answers	966	100	944	100
Missing	13		35	
Sample	979		979	

Table 3.4: Difference in education and occupational prestige between respondents and fathers

	Current		In 5 years	
Social class	No.	% of valid answers	No.	% of valid answers
Lowest (1) to 3	201	21	138	14
Class 4	405	42	361	38
Class 5	270	28	331	34
6 to highest (9)	98	10	134	14
total valid answers	974	100	964	100
missing	5		15	
sample	979		979	

Table 3.5: Self-positioning on a social distance scale, current and in 5 years

3.4.2 Respondents' choice behavior

There was a total of $979 \cdot 8 = 7,832$ decisions, of which almost 20 percent were made in favor of an alternative over the status quo (see Table 3.6). There are at least three explanations for this low percentage. First, in spite of checking in the pretests, the levels of the attributes in the experiment may not have been sufficiently extreme to make respondents switch. Second, some attributes (e.g. benefits going to the unemployed; see Table 3.8), may not have been sufficiently valued to cause a switch. Finally, there may be errors in decision making because the consistency test revealed 14 percent of choices to be inconsistent. However, there may simply be marked status quo bias in the face of highly complex decision-making situations (see the large negative constant in Table 3.8). Nonetheless, only 21 percent of respondents never opted for an alternative (see Table 3.7). Conversely, almost 80 percent departed from the status quo at least once.

Choices	No.	in percent
for alternative	1,562	19.94
for status quo	6,088	77.73
No decision	182	2.32
Total	7,832	100

Table 3.6: Total number of choices

# choices for alternative	No.	in percent
0	209	21.35
1	309	31.56
2	226	23.08
3	131	13.38
4	57	5.82
5	16	1.63
6	10	1.02
7	0	0.00
8	5	0.51
Total valid answers	965	98.57
Missing	14	1.43
Sample	979	100

Table 3.7: Distribution of the numbers of chosen alternatives per respondent

3.5 Estimation results

3.5.1 Simple model: preferences of an average respondent

Estimation of equation (3.15) includes REDIST^2 to allow for a possible nonlinearity of the indirect utility function. Moreover, it has to take into account that uses and types of beneficiaries add up to 100 percent (see Table 3.1). In order to avoid perfect collinearity, PENS (Pensioners) and OTH_FOR (Other foreigners) were dropped to obtain

$$\begin{aligned} \Delta V &= c_0 + \beta_1 \text{W\_POOR} + \beta_2 \text{UNEMP} + \beta_3 \text{ILL} + \beta_4 \text{FAM} + \\ &\quad + \gamma_1 \text{SWISS} + \gamma_2 \text{WEU\_FOR} + \\ &\quad + \delta_1 \text{REDIST} + \delta_2 \text{REDIST}^2 + \eta \text{TAX} + \varphi \end{aligned} \tag{3.16}$$

Estimation of a few of the $5 \cdot 3 = 15$ specifications with alternative exclusions produced results similar to those displayed in Table 3.8. Specifically, they agree in that additional redistribution causes respondents to opt for the alternative with a lower probability, which is even more true of an increase in the income tax to finance it [for the influence of its composition, see

Variable	Coeff.	Std. err.	z	$P > \|z\|$	Marg. eff.
Recipients' Social Group					
W_POOR	0.02784	0.00714	3.90	0.000	0.00697
UNEMP	0.01134	0.00452	2.51	0.012	0.00284
ILL	0.01600	0.00463	3.46	0.001	0.00400
FAM	0.06378	0.00942	6.77	0.000	0.01596
Recipient's Nationality					
SWISS	0.03656	0.00552	6.63	0.000	0.00915
WEU_FOR	0.02925	0.00869	3.37	0.001	0.00732
REDIST	-0.00523	0.00176	-2.97	0.003	-0.00131
REDIST^2	-0.06619	0.01174	-5.64	0.000	-0.01656
TAX	-0.02053	0.00183	-11.21	0.000	-0.00514
CONSTANT	-1.29878	0.06132	-21.18	0.000	n.a.

# observations	7,650
Log likelihood	-3,566.76
$\chi^2(0)$	108.87
Prob > χ^2	0.000
σ_u	0.41610
ρ	0.14759

Table 3.8: Random effects probit estimates for the simple model

Neustadt and Zweifel (2010a)]. Moreover, the negative constant points to a strong status quo bias. By eq. (3.14), the marginal willingness to pay (MWTP) for redistribution is given by

$$\text{MWTP}_{\text{REDIST}} = \frac{\partial \Delta V / \partial \text{REDIST}}{\partial \Delta V / \partial \text{TAX}} = -\frac{\delta_1 + 2\delta_2 \text{REDIST}}{\eta} \tag{3.17}$$

Thus, one obtains an estimated MWTP value of -0.26 percentage points of income share per additional percentage point of GDP devoted to redistribution, in excess of the status quo. Evaluated at the mean personal income of the sample, this amounts to CHF -11.78 per month. However, this figure is dwarfed by the compensation one would have to pay respondents to depart from the status quo, amounting to an estimated 63 percent of their monthly income, or 5.27 percent of their annual income.

3.5.2 Extended model: preference heterogeneity

Economic well-being and preferences for redistribution

Here, the simple model is extended by including one of the socioeconomic variables at a time (personal income, education, self-positioning on a social distance scale) as well as its interactions with the attributes. Thus, in the case of income, e.g., eq. (3.15) is modified to read[3],

$$\Delta V = c_0 + \dots + c_0' \mathrm{INC} + \dots + \beta_1 \mathrm{REDIST} + \dots + \beta_1' \mathrm{REDIST} \cdot \mathrm{INC} + \dots \tag{3.18}$$

Variable	MWTP, % of income	MWTP, CHF	Std. err., CHF	
Income group 1 (low)	-1.14215	-11.42	6.08	***
Income group 2	-0.64081	-19.22	9.37	***
Income group 3	-0.43293	-21.65	9.83	***
Income group 4 (high)	0.02117	1.81	13.47	
No high school	-0.62526	-25.13	7.12	**
High school, no college	-0.08911	-4.58	7.70	**
College	0.01501	1.04	14.71	
Social group 1 (low)	-0.40762	-14.72	8.49	***
Social group 2	-0.65405	-28.45	8.81	***
Social group 3	-0.30303	-15.06	12.36	*
Social group 4 (high)	0.25550	17.61	11.01	*

Note: *** (**,*) denotes statistical significance at the 1 (5, 10) percent level.

Table 3.9: Marginal WTP values for redistribution (in percent of monthly personal income and CHF) derived from the extended model with measures of economic well-being

Hypothesis 1 states that the demand for redistribution is expected to decrease with higher values of (a) income, (b) education, and (c) social status. Hypothesis 1(a), with its focus on personal income, cannot be confirmed (see Table 3.9). In fact, MWTP for redistribution as a percentage of income is most strongly negative in the lowest income group and consistently increases up to the second-highest. In terms of CHF amounts, negative MWTP values reach a maximum among the middle groups No. 2 and 3. However, the differences in MWTP values between Income Groups 1 and 2 ($t = 0.65$) as well as between Income Groups 2 and 3 ($t = 0.75$) are not significant. Still, differences in MWTP values within all other pairs of groups are shown to be significant at the 95 percent level (with the exception of the difference between Income Groups 1 and 3 being significant at the 90 percent level).

Similarly, Hypothesis 1(b) finds no empirical support, with MWTP values increasing rather than decreasing with higher levels of education. The evidence is mixed concerning Hypothesis

[3]The full specification is available from authors on request. The relevant results are shown in Table 3.9.

Variable	MWTP, % of income	MWTP, CHF	Std. err., CHF	Test
Downward mobility in education	-1.57572	-6.26	3.50	2a:
No mobility in education	-0.23996	-1.06	0.53	R
Upward mobility in education	-0.32110	-1.84	1.11	
Downward mobility in prestige	0.39446	1.62	1.00	2b:
No mobility in prestige	-0.38294	-1.84	1.12	(C)
Upward mobility in prestige	-0.09002	-0.51	1.22	
Downward past income mobility	-0.13457	-0.60	1.29	2c:
No past income mobility	-0.58353	-2.49	0.69	(C)
Upward past income mobility	-0.08165	-0.49	1.38	
Downward expected income mobility	0.10437	0.83	1.79	2d:
No expected income mobility	-0.55952	-2.60	0.73	(C)
Upward expected income mobility	-0.20783	-0.76	0.83	
Downward social mobility	-0.18929	-0.84	0.68	2e:
No social mobility	-0.54176	-2.52	0.75	(C)
Upward social mobility	0.14992	0.77	1.76	

Note: (C)=partially confirmed, R=rejected

Table 3.10: Marginal WTP values for redistribution (in percent of monthly personal income and CHF) derived from the extended model with mobility measures

1(c) since resistance against redistribution seems to increase from the lowest to group No. 2 of the social self-positioning scale. However, the difference between Groups 1 and 2 is only weakly significant ($t = -1.20$).

Social Mobility and Preferences for Redistribution

This time, the simple model is extended to include one of the following mobility measures: (a) intergenerational mobility in education (DIFF_ED), (b) intergenerational mobility in occupational prestige, (c) income mobility in the past, (d) expected income mobility in the future, or (e) the change in the self-positioning on a social distance scale as well as their interactions with the attributes. Therefore, in the case of the intergenerational mobility in education, eq. (3.16) is modified to become

$$\Delta V = c_0 + \dots + c_0' \text{DIFF\_ED} + \dots + \beta_1 \text{REDIST} + \dots + \beta_1' \text{REDIST} \cdot \text{DIFF\_ED} + \dots$$

Hypothesis 2 states that the demand for redistribution is expected to decrease with upward income or social mobility. In its version 2(a), it is rejected because negative MWTP is maximum among participants whose educational level is lower than their fathers', with the differences with the other two groups being highly significant (see Table 3.10). Hypothesis

2(b), with its focus on mobility in occupational prestige, finds partial support in that the MWTP of respondents with downward mobility is positive, and, the others, negative. Similarly, Hypothesis 2(c) can be accepted only to the extent that citizens with downward income mobility in the past exhibit the least resistance against redistribution. As to Hypothesis 2(d), there are weak signs suggesting that citizens with downward expected income mobility in the future might have a positive MWTP, in contrast to those with no mobility expectations. But statistical significance of two of three MWTP values is lacking to begin with, amounting to partial confirmation of Hypothesis 2(d) only. Finally, Hypothesis 2(e) is merely confirmed to the extent that individuals with downward social mobility exhibit a higher MWTP than those with no social mobility, with the corresponding t value suggesting statistical significance of the difference in MWTP values.

The one consistent pattern seems to be the following. In four out of five cases (except mobility in education), citizens with no past or future expected mobility display the highest negative MWTP values both in terms of a share in their income and in absolute amount. This seems to point to risk aversion in the face of the 'veil of ignorance' [Beck (1994)]; however, this argument has been traditionally used to predict positive rather than the observed negative MWTP for income redistribution. On the other hand, risk aversion constitutes one of the main explanations of the status quo bias (see Section 3.5.1). Neustadt and Zweifel (2010b) [see Chapter 5] conduct an analysis of the sustainability of the Swiss welfare state and show that an average respondent exhibit his maximum WTP for redistribution at the level of 21% of GDP, clearly below the status quo of 25%. Therefore, this DCE seems to suggest that Swiss citizens, while markedly risk averse, do not believe the current level of income redistribution provided by the government to be optimal for the protection against the risk impinging on their economic and social status. Such an attitude could be justifiably called realistic for citizens of a small country whose economic fortune has depended on developments abroad for decades if not centuries.

3.6 Conclusion and discussion

In this paper, we elicited citizens' willingness to pay for redistribution through a Discrete Choice experiment performed in 2008. Based on the simple model that relates choices to the attributes of redistribution only, the average Swiss citizen must be paid a compensation of

CHF 11.78 (some US$ 9.40) per month (0.02 percent of annual income) for an additional percentage point of GDP devoted to public redistribution. In addition, a very marked status quo bias would have to be overcome by payment of another 5.27 percent of annual income.

However, such an experiment also permits to test several hypotheses concerning the determinants of the demand for redistribution without any confounding supply-side influences. By including one of three measures of current economic well-being at a time, the extended model allows us to test static Hypothesis 1, stating that demand for redistribution decreases with income. However, it is found to increase with level of education and (in part) with personal income as well as higher self-positioning on a social scale.

With the inclusion of five measures of social mobility, dynamic Hypothesis 2 (POUM) could be tested as well. Except for mobility in education, citizens with no mobility at all display the highest resistance against redistribution, contrary to POUM but underscoring the importance of the status quo bias.

The analysis presented in this paper is subject to several limitations. First, only purely economic explanations of demand for redistribution (income, social mobility) were tested. However, recent contributions to the field show that up to 90 percent of cross-country differences in public spending can be related to institutional and behavioral factors [see e.g. Alesina and Glaeser (2004), Akkoyunlu et al. (2009)]. Thus, future work should be devoted to an analysis of behavioral determinants of stated willingness to pay for redistribution. A first step in this direction is made by Neustadt (2010) [see Chapter 4]. Second, the status quo bias found in this paper calls for more detailed analysis. To the extent that it reflects risk aversion, it should induce demand for redistribution - contrary to the results presented here. Finally, the evidence only relates to a point of time and thus may be subject to transitory shocks. Still, by appealing to citizens' stated preferences, the present contribution sheds some light on the debate between those who claim that there is excess redistribution and those who claim there is too little.

3.7 Acknowledgements

The authors gratefully acknowledge financial support from the Swiss National Science Foundation (SNF) under Project no. 100012-116398. They received helpful comments from Michael Brooks, Quoc Anh Do, Chris Doucouliagos, Hartmut Egger, Reinhard Madlener, Myung Jae

Sung, Arthur Schram, Alois Stutzer, Claudia Vittori, Alexey Zakharov, two anonymous referees as well as participants in the conferences 'New Directions in Welfare' (Oxford, UK, 29 June to 1 July 2009), the Australasian Public Choice Conference 2009 (Deakin University, Melbourne, Australia, 10 to 11 December 2009), the 2010 meeting of the European Public Choice Society (Izmir, Turkey, 8-11 April 2010), the 15th Spring Meeting of Young Economists (Luxembourg, 15-17 April 2010) and the 66th annual congress of the International Institute of Public Finance (Uppsala, Sweden, 23-26 August 2010) as well as in the Research Seminar at the University of Bayreuth and are grateful to Sule Akkoyunlu, Maurus Rischatsch, and Georgios Sismanidis for their assistance in the preparation of the discrete choice experiment.

3.8 Appendix

Exhibit 1: Status Quo Card (current state of redistribution)

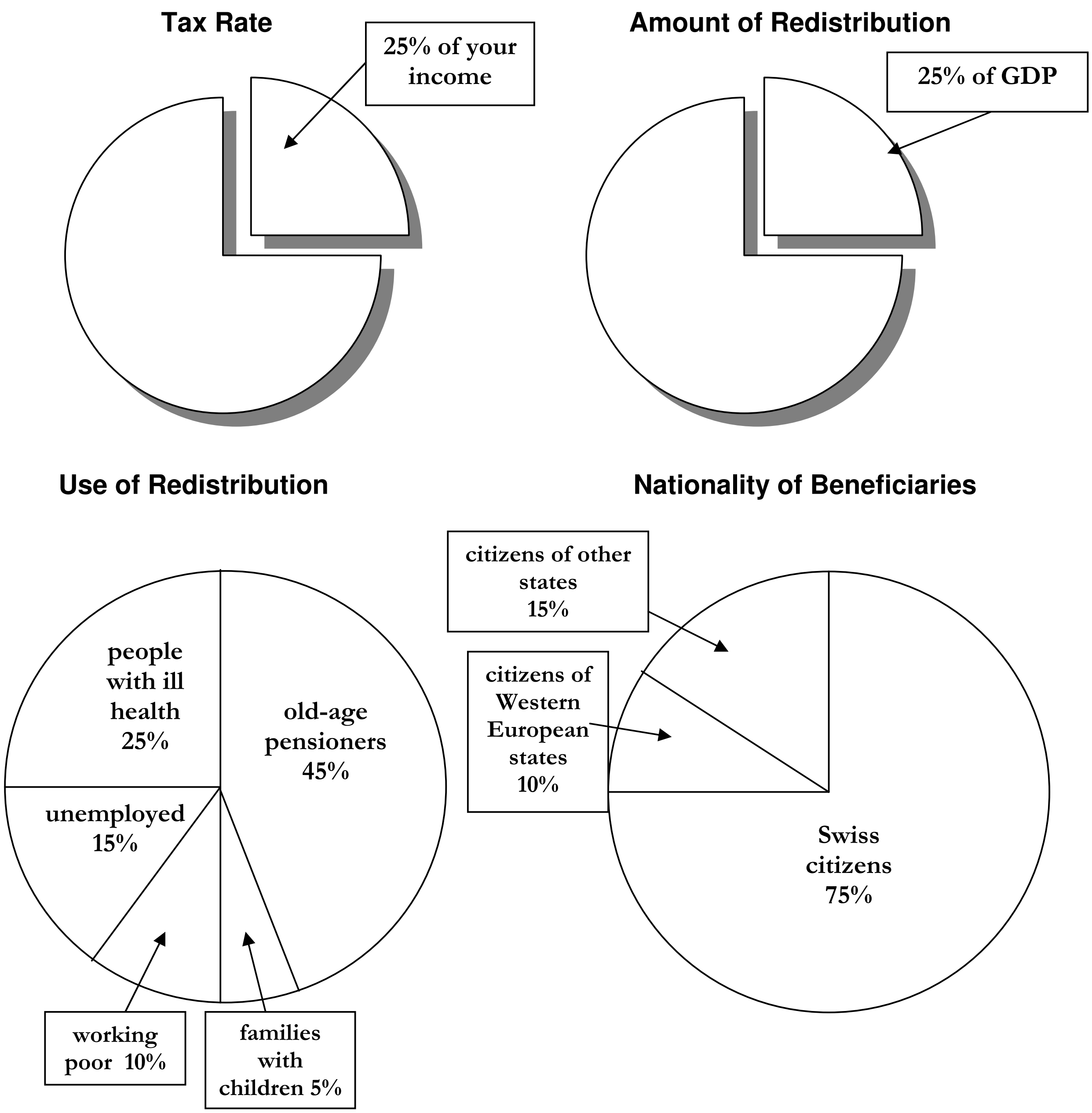


Exhibit 2: Card for Alternative No. 1

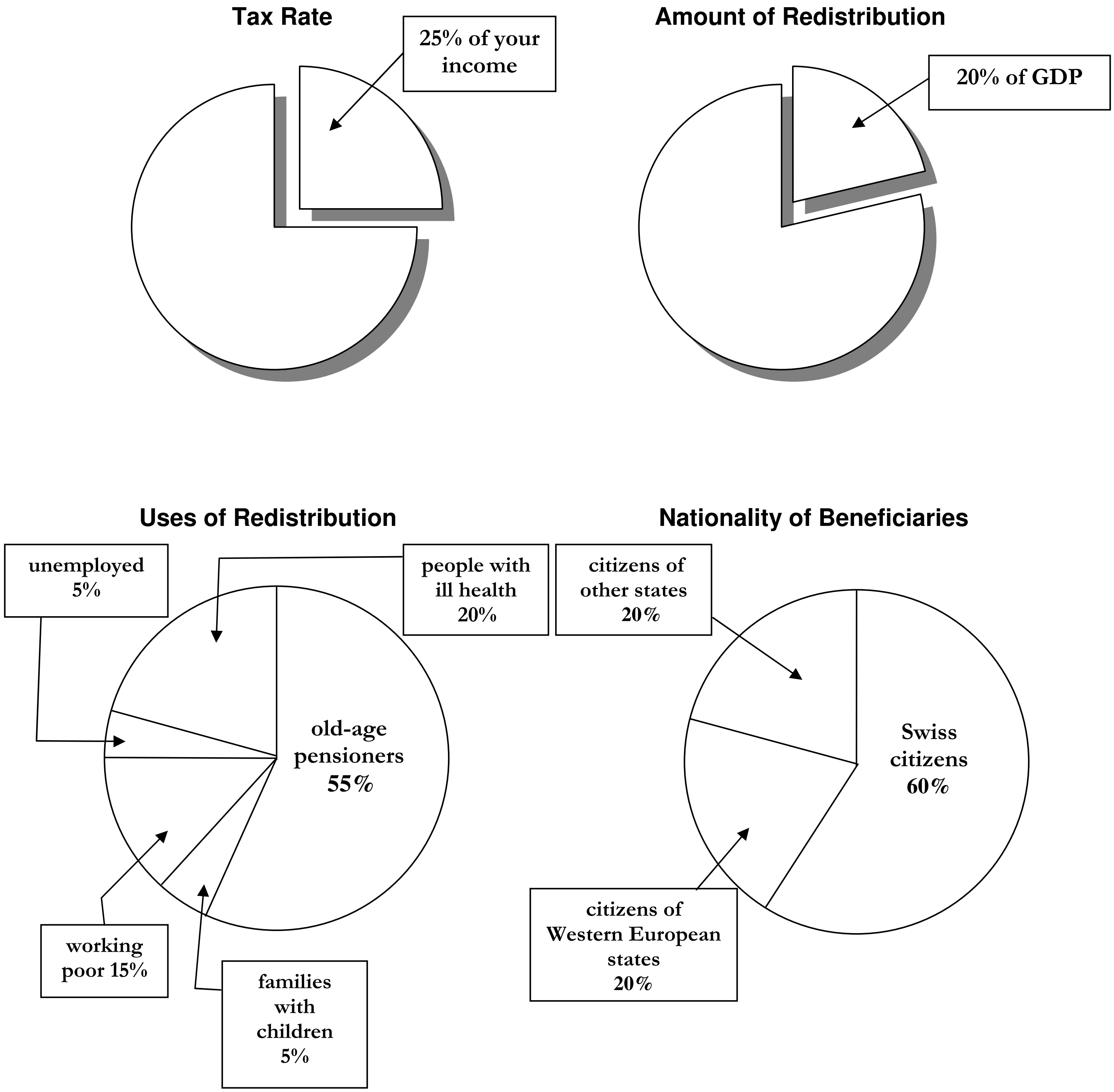
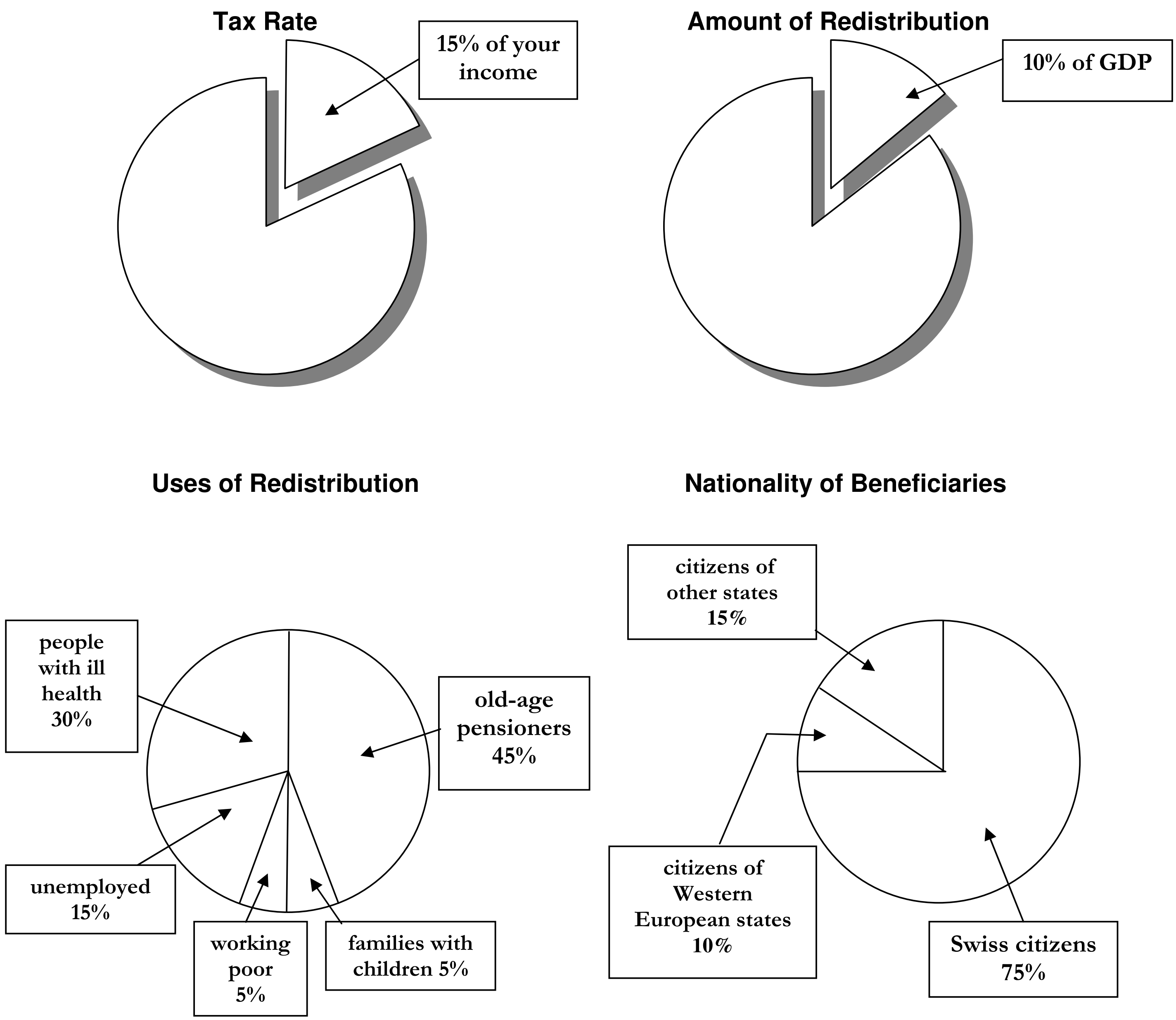


Exhibit 3: Card for Alternative No. 2



Do religious beliefs explain preferences for income redistribution? Experimental evidence

Ilja Neustadt

Submitted to “CESifo Economic Studies”.

Abstract: Due to the mixed empirical evidence bearing on the economic determinants, beliefs have been at the center of attention of research into preferences for income redistribution. We elicit preferences for income redistribution through a Discrete Choice Experiment performed in 2008 in Switzerland and relate them to several behavioral determinants, in particular to religious beliefs. Estimated marginal willingness to pay (WTP) is positive among those who do not belong to a religious denomination, and negative otherwise. However, the marginal WTP is shown to increase with a higher degree of religiosity. Moreover, those who state that luck or connections play a crucial role in determining economic success exhibit significantly higher WTP values than those who deem effort to be decisive.
Keywords: Income redistribution, beliefs, religiosity, welfare state, preferences, willingness to pay, discrete choice experiments

JEL classification: C35, C93, D63, H29

Chapter 4

Do religious beliefs explain preferences for income redistribution? Experimental evidence

4.1 Introduction

Citizens' preferences with regard to income redistribution are of crucial importance for the political debate as to the future of the welfare state. While the typical right-wing stance is to decry it as excessive, the left points to pockets of poverty even in rich societies that need to be eradicated through more redistribution. This paper contributes to this debate by measuring citizens' willingness to pay (WTP) for income redistribution through a Discrete Choice experiment (DCE) and relating it to a set of behavioral determinants. The data come from a DCE performed in the fall of 2008 and involving 979 Swiss citizens.

Recently, there has been a great deal of research into the demand for redistribution and its determinants, which will be discussed in Section 4.2 below. One line of thought relates the measured amount of redistribution to economic, institutional, and behavioral factors. Examples are Alesina and Giuliano (2009) and Akkoyunlu et al. (2009). However, the observed amount of redistribution is the outcome of an interaction between demand and supply, with supply governed by a country's political institutions and processes. This classical identification problem would have to be addressed in order to make inferences about citizens' preferences for redistribution. A second direction of research, exemplified by Alesina and La Ferrara (2005) and Guillaud (2008), relies on surveys designed to measure attitudes towards redistribution. The problem with this approach is its failure to impose a budget constraint. It therefore cannot predict actual decision making (e.g. voting at the polls), where citizens take the consequences in terms of their own income and wealth into account. A third approach seeks to solve this problem through Contingent Valuation (CV) experiments [see e.g. Boeri et al. (2001, 2002)[1]]. The weakness of the CV approach is that it holds all the attributes of the good in question constant, varying its price only. One would want to vary other attributes of redistribution besides its tax price, viz. its uses (for health, old age, etc.) and the type of beneficiary (foreigner, national).

By way of contrast, a DCE allows measurement of preferences uncontaminated by supply influences, it imposes the budget constraint through the price attribute, and it does so in a realistic way by making respondents choose between alternatives where all attributes are allowed to vary.

The remainder of this paper is structured as follows. Section 4.2 contains a literature review from which hypotheses to be tested are derived. Its first part concerns general determinants of the demand for redistribution and the second, its behavioral determinants, in particular, religious denomination, religiosity, and beliefs about the role of luck in achieving economic success. Section 4.3 presents a general description of the method of DCEs as well as the design of the present experiment. The descriptive statistics of the experiment follow in Section 4.4, and hypothesis tests, in Section 4.5. Section 4.6 summarizes the results and concludes with suggestions for future research.

[1] Boeri et al. (2001) study international attitudes towards redistribution with a focus on pension and unemployment schemes in France, Germany, Italy, and Spain. They also perform CV experiments that impose an explicit trade-off between income and social insurance coverage on respondents. They find that people oppose an extension of the welfare state, with conflicts between young and old, rich and poor, and insiders and outsiders creating significant hurdles to welfare reform.

4.2 Literature review and statement of hypotheses

This section first presents research that defines the general background of this paper and then moves on to contributions that lead to a set of specific hypotheses to be tested.

4.2.1 General determinants of the demand for income redistribution

In their reviews, Alesina and Giuliano (2009) and Akkoyunlu et al. (2009) identify a wide set of factors that can be categorized as economic, political, and behavioral determinants of the demand for income redistribution.

Economic determinants

The simplest framework for the analysis of purely economic determinants is provided by a model focusing on current economic well-being, originally proposed by Romer (1975) and Roberts (1977) and extended by Meltzer and Richard (1981) [RRMR model]. This model assumes non-altruistic utility-maximizing individuals differentiated by their income levels only. The government pays a lump-sum transfer to all citizens, financed by a linear uniform income tax. Individuals with an income below the mean favor taxation and transfers while those with an income above the mean oppose it. In a political equilibrium, the majority of voters supports a positive tax rate corresponding to the value desired by the median voter. The model's prediction is that the larger the gap between the mean and the median income, the higher the level of taxation and redistribution.

The empirical evidence is quite mixed. On the one hand, Alesina and Rodrik (1994), Persson and Tabellini (1994), and Milanovic (2000) find some supporting evidence. Furthermore, Guillaud (2008), conducting a cross-section analysis of survey data from four EU countries, shows that poorer and less educated individuals are more in favor of redistribution. On the other hand, Alesina and Glaeser (2004), Perotti (1996), and Rodriguez (1999) fail to find supporting evidence for this model. Moreover, Neustadt and Zweifel (2009) relate willingness to pay (WTP) for income redistribution elicited from a Discrete Choice Experiment (DCE, see Section 4.3.1 for details) to measures of economic well-being. WTP values are negatively related to income and education, contradicting the RRMR model.

Another economic explanation is the "Prospect of Upward Mobility" (POUM) hypothesis, suggested by Hirschman and Rothschild (1973) as the 'tunnel effect' and more recently reformulated by Benabou and Ok (2001). It extends the RRMR model by introducing individuals' expectations, based on their observations regarding the income mobility of others in society. Expected upward mobility may dampen a poor but forward-looking voter's enthusiasm for income redistribution.

Empirical support of the POUM hypothesis is provided by Alesina and La Ferrara (2005) who, using an actual mobility matrix for the United States, show that people who can expect high future income oppose redistribution. Rainer and Siedler (2008) use probabilistic expectations data to show that individuals with a sufficiently large chance of occupational upward mobility exhibit a lower demand for redistribution; conversely, those with a sufficiently large risk of occupational downward mobility opt for more redistribution. Checchi and Filippin (2004), testing the POUM hypothesis by means of a within-subjects experiment, find corroborating evidence under several alternative specifications. According to Guillaud (2008), however, individuals who subjectively experienced upward mobility over ten years tend to be more (rather than less) supportive of redistributive policies. Moreover, upward intergenerational mobility in occupational prestige goes along with more positive rather than negative attitude towards redistribution. Alesina and Giuliano (2009) examine the empirical evidence for the United States and briefly across countries, concluding that social mobility (if measured as the change in the occupational prestige) does decrease demand for redistribution once sociodemographic (age, gender, race) and socioeconomic characteristics (income, education) are controlled for. In their DCE-based study, Neustadt and Zweifel (2009) relate preferences for redistribution to mobility. They find partial empirical support for the POUM hypothesis.

Another economic explanation, suggested by the social contract literature, is that preferences for redistribution can at least in part be interpreted as a demand for insurance by risk-averse individuals. In a hypothetical situation, where individuals do not yet know their endowment as well as their future position in society ['veil of ignorance', cf. Rawls (1999)], they are predicted to exhibit positive WTP for an income transfer from more favorable future states to less favorable ones. Redistributive policies can thus be interpreted as reflecting this hypothetical demand for insurance. Beck (1994) investigates individual behavior under the 'veil of ignorance' in an experiment. By placing participants in a hypothetical society with random differences in income, represented by lotteries, he is able to derive the desired amount of income redistribution. Individuals indeed display risk aversion, albeit not of the extreme kind implied by the Rawlsian maximin rule[2]. Furthermore, they show no preference for income redistribution in excess of what can be explained by risk aversion.

Political determinants

As to the political determinants of the demand for income redistribution, the literature [Persson and Tabellini (2000, 2003); Lizzeri and Persico (2001); Milesi-Ferretti et al. (2002)] predicts that proportional representation tends towards universal programs benefitting various groups (old-age pensioners, working poor, minorities, etc.), while majority rule results in targeted 'pork barrel' programs. Persson and Tabellini (2003) find supporting empirical evidence in that countries with proportional representation have GDP shares of government expenditure that ceteris paribus are 5 percentage points higher than countries with majority rule. Moreover, Akkoyunlu et al. (2009) present weak evidence of a positive correlation between the degree of proportional representation and the transfer share in GDP in OECD countries. Additional political determinants of redistribution include two-party vs. multiparty system, presidential vs. parliamentary democracy, and direct vs. representative democracy, with two-party systems, presidential, and direct democracies all predicted to induce less public redistribution. In order to sketch the institutional background of the DCE described in Section 4.3.2, Switzerland can be described as follows. It has a high degree of proportional representation and a parliamentary democracy. Its distinguishing feature, however, is its extensive direct democratic control in the guise of popular initiatives and referenda. This might serve to limit public welfare spending while enforcing efficiency in redistribution [cf. Feld et al. (2007)].

General behavioral determinants

The mixed empirical evidence bearing on the economic determinants of preferences for redistribution calls for a detailed analysis of their behavioral determinants. In particular, beliefs have been at the center of attention. The theoretical base is laid by Alesina and Angeletos (2005), who develop a model where society's belief as to whether effort or luck determines economic success gives rise to multiple self-fulfilling equilibria. Benabou and Tirole (2006) propose a model for the emergence and persistence of such collective beliefs. Moreover, beliefs

[2]The Rawlsian maximin rule uses the maximum improvement of the individual with minimum initial wealth as the sole criterion.

can be seen as a source of altruistic preferences and inequality aversion. On the empirical side, Fong (2001) presents evidence in line with Alesina and La Ferrara (2005) suggesting that beliefs about the role of luck in determining economic success are an important determinant of the demand for redistribution. Fong (2001) also considers the effects of incentives. Fong (2001) effort determines income, then an increased income tax rate causes an output loss due to its effect on incentives. This consideration is hypothesized to qualify the link between beliefs and the demand for redistribution. However, the data fail to support this hypothesis.

While the POUM hypothesis suggests less redistribution than predicted by the RRMR model, the assumption of altruistic preferences can lead to the opposite prediction. In fact, if individuals care also about the utility of others, one might expect more redistribution than predicted by the conventional RRMR model. Fehr and Schmidt (2006) provide a review of several models of social preferences, in particular, altruism, envy, inequality aversion, fairness, and reciprocity. In a simple model of inequality aversion, it is assumed that individuals feel envy if their incomes are below that of others but they feel altruistic when their income exceeds it. Consequently, the decisive median voter demands more redistribution than in the conventional RRMR model.

Based on the assumption of inequality aversion, Neustadt and Zweifel (2010b) [see Chapter 5 of this dissertation] formulate two hypotheses to be tested. The first predicts that the citizens with higher inequality aversion exhibit a positive WTP for redistribution while those with lower inequality aversion, a negative one. The second hypothesis is based on the consideration that voters exhibiting inequality aversion tend to support the view that the government should reduce the income gap between rich and poor. Consequently, respondents who state that the reduction of the income gap is a task of the government are expected to exhibit a positive WTP for redistribution.

4.2.2 Religious beliefs and demand for income redistribution

There exists a great deal of theoretical literature dealing with religious beliefs as a determinant of demand for income redistribution, all of them predicting a negative relationship between the degree of religiosity and demand for income redistribution.

In particular, Benabou and Tirole (2006) develop a theory of collective beliefs, based on endogenous complementarities between individual cognitive choices that arise naturally from the interaction of psychological motives and economic rationality. In a simple model, they analyze an important class of religious beliefs that are linked to the 'Protestant work ethic', namely to a belief that there is a world to come, in which rewards and punishments depend on the effort and industriousness of a person during his lifetime. Alternative beliefs can be of two kinds: (i) a belief that there is no afterlife (atheism or agnosticism); (ii) a belief that there is afterlife but its rewards are not related to efforts in the current world but might be subject to observance of commandments, good deeds towards other people etc. Thus, the more religious (in the sense of the 'Protestant work ethic') a citizen is, the more effort he exerts. Thus, a more religious individual prefers lower tax rates in order to avoid income redistribution in favor of the less religious citizens who do not work as hard. If a low tax rate decided upon by a majority of religious citizens is anticipated by the population, individuals become more religious since the belief that hard work leads to rewards in afterlife generates higher utility given low income redistribution. Conversely, if a majority of citizens who happen to be less religious votes for a high level of redistribution, it can become profitable to invest in the non-religious beliefs and thus to exert less effort. In sum, two equilibria are possible:

(i) An equilibrium with a high level of religiosity in the sense of the 'Protestant work ethic', implying a high level of work effort and a low level of income redistribution.

(ii) An equilibrium with a low level of religiosity or a predominance of non-Protestant beliefs, implying a low level of work effort and a high level of income redistribution.

Moreover, Scheve and Stasavage (2006a,b) propose a model of religious participation as a substitute for insurance against adverse events. Therefore risk-averse religious individuals desire less demand for redistribution as a collective insurance device, resulting again in a negative predicted relationship between religiosity and redistribution.

A further strand of literature, based on preferences rather than beliefs, argues that public redistribution crowds out religious participation and charitable activities, giving once more rise to a negative correlation. Hungerman (2005) and Gruber and Hungerman (2007) find evidence that public insurance spending indeed crowds out religious charitable spending. Akkoyunlu et al. (2009) empirically relate the amount of public social expenditure to the strength of religious orientation for the OECD countries. They show that the partial correlation between religiosity and the share of transfers in GDP is clearly negative, supporting the theories expounded above. Moreover, Switzerland shares the somewhat guarded attitudes prevailing in the EU rather than the highly religious attitudes of the U.S. population. Being located close to the regression line, Switzerland provides additional evidence supporting the theoretical arguments relating religion to redistribution.

Based on the theoretical arguments of Benabou and Tirole (2006) and Scheve and Stasavage (2006a,b), we formulate two hypotheses to be tested in Section 4.5.2. The first predicts that members of Protestant churches exhibit a negative willingness to pay (WTP) for redistribution while members of other denominations (who do not share the Protestant work ethic but still strongly participate in and profit from private charity), a higher but still a negative one. By way of contrast, citizens with no affiliation at all who presumably share atheistic or agnostic beliefs are predicted to exhibit a strictly positive WTP for redistribution. The second hypothesis predicts the WTP for redistribution to fall with a higher level of religiosity of the individual, alternatively measured as (a) strength of the belief in God, (b) frequency of attending religious services.

HYPOTHESIS 1: *Willingness to pay for redistribution is expected to be*

(A) ***negative*** *if the individual belongs to a* ***Protestant church****,*

(B) ***negative but less so than in*** (A) *if the individual belongs to a religious denomination other than* ***Protestant****,*

(C) ***positive*** *if the individual is* ***unaffiliated****.*

HYPOTHESIS 2: *Willingness to pay for redistribution is expected to* ***decrease*** *with*

(a) *a* ***stronger*** *belief in God,*

(b) ***more frequent*** *attendance of religious services.*

4.2.3 Beliefs about the role of luck and effort and demand for income redistribution

The Benabou-Tirole model (see Section 4.2.2) suggests that beliefs about the role of effort in determining economic success or intertemporal utility are an important determinant of the preferences for redistribution. The conviction that high income and wealth are the result of work effort (belief in a just world) goes along with a low level of income redistribution. Conversely, a society that believes that luck, connections, social capital inherited from one's parents, and corruption (realistic pessimism) determine income and wealth is expected to choose a high degree of redistribution, financed by high taxes, see also Alesina and Glaeser (2004) and Alesina and Angeletos (2005). In their model of collective beliefs, Benabou and Tirole (2006) derive two possible equilibria with respect to the belief in a just world:

(i) An equilibrium with optimistic beliefs in a just world and a laissez-faire society arises in a population where the majority of citizens tries to ignore discouraging news about the efficacy of the individual effort. In turn, the majority chooses a relatively low tax rate with little redistribution and thus has strong incentives to believe that the world is indeed just.

(ii) An equilibrium with pessimistic beliefs and a welfare state arises in a population where the majority takes seriously all discouraging news about the efficacy of the individual effort. Thus, the majority chooses a high tax rate with a high level of redistribution.

Empirical evidence suggests that beliefs sharply differ between the United States and the EU. Most Americans believe that anyone can get out of poverty by hard work and that the poor remain poor only because they refuse to make the effort. By way of contrast, Europeans generally think that poverty is due to bad luck and not the individual's responsibility. Fong and Oberholzer-Gee (2007) measure the willingness-to-pay for justice in the United States using dictator games. Dictators were given $10 to split between themselves and recipients. The authors find that one third of the dictators are willing to pay one dollar out of ten for obtaining the information whether poverty was due to disability or substance abuse. Finally, Alesina and Giuliano (2009) show that a history of misfortune in the recent past such as unemployment and personal trauma makes people more risk-averse and less optimistic about upward mobility. These changes in beliefs are found to have a positive and significant effect on redistribution.

Therefore, the Benabou-Tirole model suggests that equilibrium (i) with a laissez-faire society is likely to persist in the United States while equilibrium (ii) with a full-fledged welfare state is sustainable in Europe. Akkoyunlu et al. (2009) relate the amount of public social spending to a score that ranges from 1 (hard work always brings a better life) to 10 (hard work does not bring any success) using data from OECD countries. The U.S. score is closest to 1 but lies below the regression line, while Germany and Denmark mark the other extreme. Again, as in the case of religiosity, this regression is one of two best-fitting bivariate regressions designed to explain the share of transfers in GDP. Here again, Switzerland as a test case lies right near the regression line, lending additional support to the hypothesis.

Based on the presented literature review we state Hypothesis 3 to be tested in Section 4.5.2 as follows.

HYPOTHESIS 3: *Willingness to pay for redistribution is expected to* ***increase*** *with a* ***stronger*** *individual belief that luck rather than effort determine economic success.*

4.3 Discrete choice experiments

4.3.1 Theoretical foundations

Discrete Choice Experiments (DCEs) provide a tool for measuring individuals' preferences for characteristics of commodities, the so-called attributes. In contradistinction to classical Revealed Preference Theory, originating with Samuelson (1938), DCEs allow individuals to express their preferences for non-marketed as well as hypothetical products. During a DCE, respondents are repeatedly asked to compare the status quo with several hypothetical alternatives defined by their attributes including a price. By varying the levels of attributes, different product alternatives are generated. A rational individual will always choose the alternative with the highest utility. From the observed choices, the researcher can infer the utility associated with the attributes. The proposed method, derived from the New Demand Theory of Lancaster (1971), is also known as Conjoint Analysis [Louviere et al. (2000)].

The most prominent alternative to a DCE is Contingent Valuation (CV). A certain situation or product is described in detail, and respondents are asked to indicate their maximum willingness to pay (WTP) for this fixed product. Only its price attribute is varied, while in Conjoint Analysis all relevant attributes are varied simultaneously, making it a multi-attribute valuation method [Merino-Castello (2003)]. While a DCE describes the product in less detail than a typical CV study, it allows for analyzing many product varieties by varying the levels of relevant attributes [Louviere et al. (2000), p. 344]. Trade-offs among attributes can be explicitly taken into account and WTP values of attributes estimated separately (see below).

Furthermore, strategic behavior of respondents is less likely than in CV with its exclusive emphasis on price, which facilitates strategic behavior. Finally, biases that easily occur when individuals are directly asked about their WTP are less frequently observed in a DCE [Ryan (2004)].

A particular advantage of a DCE in the present context is that it permits to explicitly impose the budget constraint through a price attribute in the guise of the tax share of income used to finance the transfers considered. Respondents can be made to simultaneously choose this share and hence the 'size of the pie' and the 'slices of the pie' devoted to different types of recipients and uses (health, old age, etc.; see Exhibits No. 1 to 3 in Appendix). Thus, trade-offs among different attributes of the redistribution plan can be calculated to assess the relative importance of the respective redistributive goals.

The econometric method used is based on the Random Utility Theory [see Luce (1959), Manski and Lerman (1977) and McFadden (1974, 1981, 2001)]. Individual i values alternative j according to the utility V_{ij} attained, which is given by

$$V_{ij} = v_i(a_j, p_j, y_i, s_i, \varepsilon_{ij}). \tag{4.1}$$

Here, $v_i(\cdot)$ denotes i's indirect utility function, a_j, the amount of attributes associated with alternative j, and p_j, price. The individual's income and sociodemographic characteristics are symbolized by y_i and s_i, respectively. Finally, ε_{ij} denotes the error term, which is due to the fact that the experimenter will never observe all the arguments entering v_i, imparting a stochastic element to observed choices. As usual, the utility function is additively split into a systematic component $w(\cdot)$ and a stochastic one,

$$V_{ij} = w_i(a_j, p_j, y_i, s_i) + \varepsilon_{ij}.$$

A utility-maximizing individual i will prefer alternative j to alternative l if and only if

$$w_i(a_l, p_l, y_i, s_i) + \varepsilon_{il} \leq w_i(a_j, p_j, y_i, s_i) + \varepsilon_{ij}. \tag{4.2}$$

Due to the presence of the stochastic term, only the probability P_{ij} of individual i choosing alternative j rather than alternative l can be estimated, with

$$\begin{aligned} P_{ij} &= \text{Prob}\left[w_i(a_l, p_l, y_i, s_i) + \varepsilon_{il} \leq w_i(a_j, p_j, y_i, s_i) + \varepsilon_{ij}\right] \qquad (4.3) \\ &= \text{Prob}\left[\varepsilon_{il} - \varepsilon_{ij} \leq w_i(a_j, p_j, y_i, s_i) - w_i(a_l, p_l, y_i, s_i)\right]. \qquad (4.4) \end{aligned}$$

Thus, the probability of choosing j amounts to the probability of the systematic utility difference $w_i[j]-w_i[l]$ dominating the 'noise', $\varepsilon_{il}-\varepsilon_{ij}$. The error terms $\{\varepsilon_{il}, \varepsilon_{ij}\}$ can be assumed to be normally distributed with mean zero and variances σ_l^2 and σ_j^2 as well as covariance σ_{lj}. Under these assumptions, $\varphi_{ij} := \varepsilon_{il} - \varepsilon_{ij}$ is also normally distributed with mean zero and variance $\sigma^2 := \text{Var}[\varphi_{ij}] = \sigma_l^2 + \sigma_j^2 - 2\sigma_{lj}$. Thus, equation (4.4) can be represented as

$$P_{ij} = \Phi\left(\frac{w_i(a_j, p_j, y_i, s_i) - w_i(a_l, p_l, y_i, s_i)}{\sigma}\right), \qquad (4.5)$$

where $\Phi(\cdot)$ denotes the cdf of a standard normal distribution. This model is known as the binary probit model [cf. Ben-Akiva and Lerman (1985)]. Hensher et al. (1999) provide empirical evidence that a linear specification of the function $w(\cdot)$ leads to good predictions in its middle ranges. Therefore, in the case of the simple model that relates utilities and choice probabilities to the attributes only (see Section 4.5.1), one posits

$$w_i(a_j, p_j, y_i, s_i) = c_i + \sum_{k=1}^{K} \beta_k a_{kj} + \varepsilon_{ij}, \qquad (4.6)$$

where c_i represents an individual-specific constant, a_k, $k = 1, \ldots, K$, are the attributes of the alternative, and β_k, $k = 1, \ldots, K$, are the parameters to be estimated. These parameters can be interpreted as the constant marginal utilities of the corresponding attributes. One obtains the following expression representing the difference in utility of individual i between alternative j and status quo,

$$\Delta V_{ij} = c_i + \sum_{k=1}^{K} \beta_k \Delta a_{kj} + \beta_p \Delta p_j + \varphi_{ij}, \qquad (4.7)$$

where $\Delta a_{kj} = a_{kj} - a_{lj}$, $\Delta p_j = p_j - p_l$, $c_i = c_{il} - c_{ij}$, and $\varphi_{ij} = \varepsilon_{il} - \varepsilon_{ij}$ for each $j \neq l$. The marginal rate of substitution between two attributes m and n is given by

$$\text{MRS}_{m,n} = -\frac{\partial v / \partial a_m}{\partial v / \partial a_n}. \tag{4.8}$$

Therefore, the marginal WTP for attribute a_m can be calculated by dividing the marginal utility of this attribute by the marginal utility of the price attribute [in the present context, the income tax rate, see e.g. Telser (2002), p. 56]:

$$\text{MWTP}(a_m) = \frac{\partial v / \partial a_m}{\partial v / \partial p_j}. \tag{4.9}$$

For econometric inference, it is important to recall that the same individual makes several choices. The two-way random-effect specification takes this into account with $\varphi_{ij} = \mu_i + \eta_{ij}$, where μ_i denotes the component that varies only across individuals but not across the choice alternatives. The terms μ_i and η_{ij} are assumed uncorrelated with the product attributes $(a_{i1}, \ldots, a_{iK})$ and between themselves. By a standard assumption in a probit model, $\sigma_\eta = 1$. Hence $\text{Var}[\varphi_{ij}] = \sigma_\eta^2 + \sigma_\mu^2 = 1 + \sigma_\mu^2$ and $\text{Corr}[\varphi_{ij}, \varphi_{il}] = \frac{\sigma_\mu^2}{1+\sigma_\mu^2} =: \rho$. The parameter ρ indicates how strongly the various responses of an individual are correlated with each other, or, equivalently, the share of the total variance that can be explained by individual-specific error term. The random-effects specification is justified if ρ is high and significant.

The simple model can be extended by including various socioeconomic variables (e.g. income group, level of education, social mobility). These variables need to be interacted with the product attributes as well as with the constant, giving rise to the extended model specification which allows to check for preference heterogeneity and thus to test Hypotheses 1 to 3 in Section 4.5.2. By means of a t test we can investigate whether the differences in marginal WTP values between different socioeconomic groups are statistically significant. The computation of the variance of the marginal WTP values is performed by the delta method, cf. Hole (2007)

4.3.2 Experimental design

A representative telephone survey with 979 respondents was conducted in the fall of 2008. Prior to the survey, the attributes and their levels used to define 'income redistribution' had

Attribute	Label	Status Quo Level	Alternative Levels
Shares of benefits going to			
• Working Poor	W_POOR	10%	5%, 15%
• Unemployed	UNEMP	15%	5%, 25%
• Old-Age Pensioners	PENS	45%	35%, 55%
• Families with Children	FAM	5%	10%
• People with Ill Health	ILL	25%	20%, 30%
Shares of benefits going to			
• Swiss citizens	SWISS	75%	60%, 85%
• Western European foreigners	WEU_FOR	10%	5%, 20%
• Other foreigners	OTH_FOR	15%	10%, 20%
Total amount of redistribution	REDIST	25% (of GDP)	10%, 20%, 30%, 40%, 50%
Income tax	TAX	25% (of personal income)	10%, 15%, 40%

Table 4.1: Attributes and their levels

been checked in two pretests for their relevance. Attributes form four groups (see Table 4.1).

1. Shares of the total redistribution budget to be spent on five types of recipients (viz. the working poor, the unemployed, old-age pensioners, families with children, and people with ill health);

2. Shares of the total redistribution budget to be spent on three groups (viz. Swiss citizens, western European foreigners, and other foreigners);

3. Total amount of redistribution, defined as a share of GDP;

4. Personal income tax rate to be paid by the respondent (the price attribute).

Clearly, these attributes and their levels combine to form a total number of possible scenarios that cannot be realized in an experiment. The scenarios define the n rows of the observation matrix X, with associated covariance matrix $\Omega = \sigma^2 \left(X'X\right)^{-1}$ of parameters β to be estimated. So-called D-efficient design calls for the minimization of the geometric mean of the eigenvalues of Ω,

$$D \text{ efficiency} = \left(|\Omega|^{\frac{1}{K}}\right)^{-1}$$

where K denotes the number of parameters to be estimated [cf. Carlsson and Martinsson (2003)]. Using this optimization procedure and incorporating several restrictions, the number of alternatives was reduced to 35 and randomly split into five groups. One alternative was included twice in each decision set for a consistency test, resulting in 8 binary choices per respondent.

In order to make sure that decisions were based on a homogeneous information set and made in a consistent way, respondents were provided with a detailed description of the attributes and their possible realizations. The Appendix shows the graphical representation of the status quo (Exhibit 1) and two selected alternatives (Exhibits 2 and 3).

4.4 Descriptive statistics

4.4.1 Socioeconomic characteristics

The sample consists of 979 respondents, 70 percent of them residing in the German-speaking part and 30 percent in the French-speaking part of Switzerland. Some 94 percent were born in the country, 50 percent are men, 20 percent having a monthly income below CHF 2,000 and 23 percent, above CHF 6,000, reflecting the structure of the Swiss population. However, only 1.5 percent of the respondents are unemployed.

39 percent of the respondents are members of the Roman Catholic Church while 51 percent belong to the Reformed Church[3]. An additional 2 percent are members of other religious denominations while 8 percent are not affiliated [see Table 4.2]. As to the strength of religious beliefs, 39 percent indicated no or weak belief in God as well as moderate belief while 22 percent of respondents claimed to have a strong or a very strong belief in God [see Table 4.3]. Moreover, 27 percent of respondents attended a religious service at least once in the last month. Individuals' shares whose last service attendance was 1 to 2 months ago and 3 to 6 months ago made up 27 and 22 percent, respectively [see Table 4.4]. Finally, 24 percent stated not having attended a service within the last 6 months. However, the number of missing answers with 109 is unusually high, probably due to the fact that many individuals who never attended a religious service preferred to refuse their answers.

Table 4.5 shows the distribution of answers to the question, "Is work effort or luck and connections more important for economic success?", with step 1 indicating the belief that work effort alone determines success, and step 10, the opposite belief that work effort does not matter at all. The majority of respondents seem to believe in the role of effort. In fact, 25 percent of respondents placed themselves on steps 1 or 2, 23 percent, on step 3, and 16

[3] Largest Protestant denomination in Switzerland.

Religious denomination	No.	% of valid answers
Roman Catholic Church	383	39
Reformed Church	494	51
Unaffiliated	80	8
Other	19	2
Total valid answers	976	100
Missing	3	
Sample	979	

Table 4.2: Religious denomination of the respondents

strength of the belief	No.	% of valid answers
no or weak belief	382	39
moderate belief	384	39
strong or very strong belief	209	22
total valid answers	975	100
missing	4	
sample	979	

Table 4.3: Level of respondents' religiosity measured as strength of their belief in God

percent, on step 4. As much as 24 percent chose step 5 while only 12 percent placed themselves on steps 6 to 10.

4.4.2 Respondents' choice behavior

There is a total of $979 \cdot 8 = 7,832$ decisions, of which not quite 20 percent were made in favor of an alternative over the status quo [see Table 4.6]. There are at least three explanations for this low percentage. First, in spite of checking in the pretests, the levels of the attributes in the experiment may not have been sufficiently spaced apart to make respondents switch. Second, some attributes (e.g. benefits going to the unemployed; see Table 4.8), may not have been important enough to cause a switch. Finally, there may be errors in decision making because the consistency test revealed 14 percent of choices to be inconsistent. However, there may simply be marked status quo bias in the face of highly complex decision-making situations, as suggested by the large negative constant in Table 4.8. Nonetheless, only 21 percent of respondents never opted for an alternative [see Table 4.6]. Conversely, almost 80 percent departed from the status quo at least once.

Last attendance	No.	% of valid answers
less than 1 month ago	236	27
1 to 2 months ago	193	22
3 to 6 months ago	236	27
more than 6 months ago	205	24
total valid answers	870	100
Missing	109	
Sample	979	

Table 4.4: Level of respondents' religiosity measured by time of their last attendance of a religious service

	No.	% of valid answers
steps 1 to 2	247	25
step 3	226	23
steps 4 to 5	389	40
steps 6 to 10	112	12
total valid answers	974	100
Missing	5	
Sample	979	

Table 4.5: Belief whether effort or luck determine economic success on a scale from 1 to 10. Step 1 indicates the belief that only effort determines success, step 10 indicates the belief that only luck determines success.

4.5 Estimation results

4.5.1 Simple model: preferences of an average respondent

Estimation of equation (4.7) includes REDIST^2 to allow for a possible nonlinearity of the indirect utility function with regard to the GDP share of redistribution REDIST. Moreover, the fact that uses and types of beneficiaries add up to 100 percent needs to be taken into account [see Table 4.1]. In order to avoid perfect collinearity, PENS (Pensioners) and OTH_FOR (Other foreigners) were dropped to obtain

$$\begin{aligned} \Delta V &= c_0 + \beta_1 \text{W\_POOR} + \beta_2 \text{UNEMP} + \beta_3 \text{ILL} + \beta_4 \text{FAM} + \\ &\quad + \gamma_1 \text{SWISS} + \gamma_2 \text{WEU\_FOR} + \\ &\quad + \delta_1 \text{REDIST} + \delta_2 \text{REDIST}^2 + \eta \text{TAX} + \varphi \end{aligned} \tag{4.10}$$

Estimation of a few of the $5 \cdot 3 = 15$ specifications with alternative exclusions produced

Choices	No.	in percent
for alternative	1,562	19.94
for status quo	6,088	77.73
No decision	182	2.32
Total	7,832	100

Table 4.6: Total number of choices

# choices for alternative	No.	in percent
0	209	21.35
1	309	31.56
2	226	23.08
3	131	13.38
4	57	5.82
5	16	1.63
6	10	1.02
7	0	0.00
8	5	0.51
Total valid answers	965	98.57
Missing	14	1.43
Sample	979	100

Table 4.7: Distribution of the number of chosen alternatives per respondent

results similar to those displayed in Table 4.8. Specifically, they agree in that alternatives with additional redistribution are chosen with a lower probability [for details with regard to 'slices' of the pie, see Neustadt and Zweifel (2010a)]. Also, note the sizeable and highly significant coefficient of the price attribute TAX, which is important for the estimation of marginal willingness-to-pay (MWTP) values [see eq. (4.9)]. For redistribution, the MWTP value is given by

$$\text{MWTP}_{\text{REDIST}} = \frac{\partial \Delta V / \partial \text{REDIST}}{\partial \Delta V / \partial \text{TAX}} = -\frac{\delta_1 + 2\delta_2 \text{REDIST}}{\eta} \tag{4.11}$$

This amounts to -0.25 percentage points of income share per additional percentage point of GDP devoted to redistribution in excess of the status quo. Evaluated at the mean personal income of the sample, this equals CHF -11.78 per month. However, this figure is dwarfed by the compensation one would have to pay respondents to depart from the status quo,

Variable	Coeff.	Std. err.	z	$P > \|z\|$	Marg. eff.
Recipients' Social Group					
W_POOR	0.02784	0.00714	3.90	0.000	0.00697
UNEMP	0.01134	0.00452	2.51	0.012	0.00284
ILL	0.01600	0.00463	3.46	0.001	0.00400
FAM	0.06378	0.00942	6.77	0.000	0.01596
Recipient's Nationality					
SWISS	0.03656	0.00552	6.63	0.000	0.00915
WEU_FOR	0.02925	0.00869	3.37	0.001	0.00732
REDIST	-0.00523	0.00176	-2.97	0.003	-0.00131
REDIST^2	-0.06619	0.01174	-5.64	0.000	-0.01656
TAX	-0.02053	0.00183	-11.21	0.000	-0.00514
CONSTANT	-1.29878	0.06132	-21.18	0.000	n.a.

# observations	7,650
Log likelihood	-3,566.76
$\chi^2(0)$	108.87
Prob > χ^2	0.000
σ_u	0.41610
ρ	0.14759

Table 4.8: Random effects probit estimates for the simple model

amounting to an estimated 63 percent of their monthly income, or 5.27 percent of their annual income [see the large negative constant in Table 4.8].

Neustadt and Zweifel (2010b) [see Chapter 5] construct the (quadratic) WTP function and show that it attains a maximum at 21.05% of GDP, definitely below the current value of 25%. Therefore, they argue that Swiss welfare state is too big in the light of average citizens' preferences.

4.5.2 Extended model: preference heterogeneity

Religious denomination and preferences for redistribution

The simple model is now extended by including dummies for the religious denomination [see Table 4.2]. The four levels of this variable are represented by three dummy variables, REF, CATH, and OTH_DEN. For instance, the former is defined as

$$\text{CATH} = \begin{cases} 1 & \text{if the respondent belongs to the Catholic Church,} \\ 0 & \text{otherwise.} \end{cases}$$

The reference category is UNAFF (unaffiliated), indicating that the respondent does not belong to a religious denomination. Since an attribute's marginal utility may vary with religious denomination, eq. (4.10) is modified to also contain interaction terms involving the denomination variables, resulting in

$$\begin{aligned}\Delta V' &= c_0' + \cdots + \alpha_1'\text{CATH} + \cdots + \alpha_2'\text{REDIST} + +\alpha_3'\text{REDIST}^2 + \ldots \\ &\quad +\lambda_2'\text{REDIST}\cdot\text{CATH} + \lambda_3'\text{REDIST}^2\cdot\text{CATH} + \ldots \\ &\quad +\lambda_4'\text{REDIST}\cdot\text{REF} + \lambda_5'\text{REDIST}^2\cdot\text{REF} + \\ &\quad +\lambda_6'\text{REDIST}\cdot\text{OTH\_DEN} + \lambda_7'\text{REDIST}^2\cdot\text{OTH\_DEN} + \cdots + \varphi'.\end{aligned}$$

	exp. sign	MWTP, % of income	MWTP, CHF	s.e., CHF	
Catholics	-	-0.32857	-15.44	7.34	***
Reformed	-	-0.47866	-21.33	11.24	***
Unaffiliated	+	0.71988	37.77	9.77	***
Others		-1.15630	-49.30	86.45	

Note: *** denotes statistical significance of MWTP in % of income at the 1 percent level.

Table 4.9: Marginal WTP values for redistribution (in percent of monthly personal income and CHF) derived from the extended model with religious denominations

Hypothesis 1(A) states that the demand for redistribution is expected to be negative for the respondents belonging to the Reformed Church. It is confirmed, with the MWTP for one percentage point increase of the total amount of redistribution being a negative CHF -15.44 [see Table 4.9]. Hypothesis 1(B), stating that Catholics exhibit a negative demand for redistribution that is, however, higher than that of the Protestants is confirmed, too. However, a t test shows that the difference in MWTP values between these two religious groups is statistically not significant. Further, Hypothesis 1(C), predicting the demand for redistribution of unaffiliated citizens to be positive finds strong empirical support with the corresponding MWTP of CHF 37.77 for one percentage point increase of the total amount of redistribution. Moreover, t tests confirm that MWTP values of the unaffiliated individuals on the one side and Catholics or Reformed on the other side are significantly different (with t values of 4.35 and 3.97, respectively).

Religiosity and preferences for redistribution

In this section, the simple model is extended by one of two measures of the level of religious participation by including the corresponding dummies [see Tables 4.3, 4.4]. For instance, in the case of the strength of religious beliefs, the three levels of this variable are represented by two dummy variables, namely WEAK (no or weak belief in God) and STRONG (strong or very strong belief), with the former being defined as

$$\text{WEAK} = \begin{cases} 1 & \text{if the respondent has no or weak belief in God,} \\ 0 & \text{otherwise.} \end{cases}$$

Here, the reference category is MODER, indicating that the respondent stated having beliefs of moderate strength. Since an attribute's marginal utility may vary with religious denomination, eq. (4.10) is modified to also contain interaction terms involving the denomination variables, resulting in

$$\begin{aligned} \Delta V'' &= c_0'' + \cdots + \alpha_1''\text{WEAK} + \cdots + \alpha_2''\text{REDIST} + \\ &\quad + \alpha_3''\text{REDIST}^2 + \cdots + \kappa_2''\text{REDIST} \cdot \text{WEAK} + \\ &\quad + \kappa_3''\text{REDIST}^2 \cdot \text{WEAK} + \cdots + \\ &\quad + \kappa_4''\text{REDIST} \cdot \text{STRONG} + \\ &\quad + \kappa_5''\text{REDIST}^2 \cdot \text{STRONG} + \cdots + \varphi'' \end{aligned}$$

	MWTP, % of income	MWTP, CHF	s.e., CHF	
no or weak belief	-0.47477	-20.67	12.75	**
moderate belief	-0.42066	-19.33	7.93	***
strong belief	0.24983	12.83	8.56	*

Note: *** (**,*) denotes statistical significance of MWTP in % of income at the 1 (5, 10) percent level.

Table 4.10: Marginal WTP values for redistribution (in percent of monthly personal income and CHF) derived from the extended model with strength of religious beliefs

Hypothesis 2(a) with its focus on the degree of religiosity as a determinant of WTP for redistribution states that the demand for redistribution is expected to decrease with a stronger belief in God. The estimated WTP values suggest to reject this hypothesis, however. In fact, the WTP increases with the strength of religious beliefs [see Table 4.10]. The t test for preference heterogeneity confirms that the WTP of individuals with strong beliefs significantly differs from the WTP of the other two groups. However, the difference between respondents with weak and moderate beliefs, respectively, cannot be shown to be statistically significant.

	MWTP, % of income	MWTP, CHF	s.e., CHF	
less than 1 month ago (group 1)	-0.90140	-38.47	38.43	*
1 to 2 months ago (group 2)	-0.35836	-15.54	12.52	*
3 to 6 months ago (group 3)	-0.44118	-20.64	10.02	***
more than 6 months ago (group 4)	-0.05989	-2.95	7.38	

Note: *** (**,*) denotes statistical significance of MWTP in % of income at the 1 (5, 10) percent level.

Table 4.11: Marginal WTP values for redistribution (in percent of monthly personal income and CHF) derived from the extended model with frequency of attending a religious service

Hypothesis 2(b) states that WTP values are predicted to decrease with a higher frequency of attendance of religious services. In fact, the estimated WTP values [see Table 4.11] seem to confirm this hypothesis. However, as indicated by the test for heterogeneity, the confidence intervals of the estimated WTP values overlap, with the notable exception of group 3 and group 4 exhibiting a weakly significant difference.

Beliefs about the role of luck and effort and preferences for redistribution

Next, the simple model is extended by including the dummy variables describing the respondents' beliefs about the role of effort vs. luck for achieving economic success. There are five corresponding dummy variables defined as follows: LUCK12 (=1 if the respondent placed himself on steps 1 or 2 on a scale from 1 to 10, =0 otherwise), LUCK3, (=1 if the respondent placed himself on step 3, =0 otherwise[4]), and LUCK6+ (=1 if the respondent placed himself on steps 6 to 10, thereby indicating a strong belief in luck, =0 otherwise) as well as their

[4]For the distribution of answers, see Table 4.5.

interactions with the attributes. The reference category is LUCK45 (=1 if the respondent placed himself on steps 4 or 5, =0 otherwise). Thus, eq. (4.10) is modified to read,

$$
\begin{aligned}
\Delta V''' &= c_0''' + \cdots + \alpha_1'''\text{LUCK12} + \cdots + \alpha_2'''\text{REDIST} + \\
&\quad + \alpha_3'''\text{REDIST}^2 + \cdots + \kappa_2'''\text{REDIST} \cdot \text{LUCK12} + \\
&\quad + \kappa_3'''\text{REDIST}^2 \cdot \text{LUCK12} + \cdots + \varphi'''
\end{aligned}
$$

	MWTP, % of income	MWTP, CHF	s.e., CHF	
steps 1 or 2 (group I)	-0.74183	-32.92	13.25	***
step 3 (group II)	-0.59576	-27.23	13.63	***
steps 4 or 5 (group III)	-0.06592	-3.08	8.12	
steps 6 to 10 (group IV)	0.71922	35.15	11.83	***

Note: *** (**,*) denotes statistical significance of MWTP in % of income at the 1 (5, 10) percent level.

Table 4.12: Marginal WTP values for redistribution (in percent of monthly personal income and CHF) derived from the extended model with beliefs as to whether effort or luck determine economic success. A higher step indicates a weaker belief in a just world and a stronger belief in luck.

Hypothesis 3 states that the demand for redistribution is predicted to increase with a stronger belief that luck determines economic success. Therefore, respondents who chose higher steps on a scale from 1 to 10 are expected to exhibit higher WTP values. As indicated by the estimation results presented in Table 4.12, this hypothesis is confirmed. In particular, individuals who placed themselves on steps 6 to 10, thereby indicating that they deem luck and connections to be crucial in the determination of income and wealth, exhibit a strongly positive marginal WTP value of CHF 35.15 for a 1 percentage point increase in the amount of redistribution as a share of GDP. The corresponding MWTP values for other groups, while being negative, do increase with higher steps. Furthermore, the t test results confirm that all differences between the MWTP values of the four groups are significant, the only exception being the difference between group I (steps 1 or 2) and group II (step 3), both believing in a just world.

4.6 Conclusion and discussion

In this paper, we elicited citizens' willingness to pay (WTP) for redistribution through a Discrete Choice experiment performed in 2008. Based on a simple model that relates choices to the attributes of redistribution only, the average Swiss citizen would have to be paid a compensation of CHF 11.78 (some US$ 9.40) per month (0.25 percent of monthly income) for an additional percentage point of GDP devoted to public redistribution. In addition, a very marked status quo bias would have to be overcome by payment of another 63 percent of monthly income.

Furthermore, we tested several hypotheses concerning the behavioral determinants of the demand for redistribution without any confounding supply-side influences. In particular, Hypothesis 1 states that it is negative among church members and positive among those without religious affiliation. An extended model that includes the pertinent variable indicating religious denomination as a regressor yields confirming evidence for this statement; however, the additional prediction that Protestants exhibit a lower WTP than Catholics finds only partial support with the respective difference between the WTP values being statistically not significant. Hypothesis 2 predicts that more religious citizens who are more likely to engage in private charity and frequently consider religion as a means of insurance (crowding-out effect) do not demand a high level of public redistribution. Here, the extended version of the model (in both alternative versions) does not support the hypothesis. Finally, Hypothesis 3 predicts that citizens with a strong belief that luck rather than effort determine economic success exhibit a higher WTP. The corresponding extended model that includes the stated belief about the role of luck as a determinant of income and wealth empirically confirms this hypothesis. In fact, Hypothesis 3 is the most successful one in predicting the citizens' demand for redistribution, providing corroborating evidence for the theoretical model of collective beliefs developed by Benabou and Tirole (2006).

The analysis presented in this paper is subject to several limitations. First, only some behavioral explanations of the demand for redistribution were tested while others (risk aversion, inequality aversion) were not controlled for. Furthermore, as suggested by recent contributions to literature in the field of public choice, citizens' preferences can be importantly influenced by political institutions, in particular by party programs [see e.g. Schläpfer et al. (2007)]. Thus, future work should be devoted to a detailed analysis of political party preferences in order to find out whether these factors also influence stated WTP for redistribution. This analysis would, however, require addressing the identification problem once again, since the supply of public redistribution is governed by political institutions. Second, the status quo bias found in this paper calls for more detailed analysis. To the extent that it reflects risk aversion, it should induce demand for redistribution - contrary to the results presented here. One possible explanation of this phenomenon can be the fact that there are some preferences that are not fully formed [see e.g. Stutzer et al. (2007)]. Another possible explanation might be the redistribution illusion, namely the fact that some respondents are not aware of the actual status quo. Finally, the evidence only relates to a point of time in one country and thus may be subject to transitory shocks and country-specific influences. Still, by appealing to citizens' stated preferences, the present contribution sheds some light on the question whether religious and cultural beliefs can explain preferences for income redistribution.

4.7 Acknowledgements

The author gratefully acknowledges financial support from the Swiss National Science Foundation (SNF) under Project no. 100012-116398. He received helpful comments from Martin Beckmann, Douglas Bernheim, Joan Costa-i-Font, Mireia Jofre-Bonet, Reinhard Madlener, Erik Schokkaert and Peter Zweifel as well as from participants in the Workshop on Behavioural Welfare Economics (Venice Summer Institute, Venice, Italy, 21-22 July 2010), organized by Joan Costa-i-Font and Frank Cowell.

4.8 Appendix

Exhibit 1: Status Quo Card (current state of redistribution)

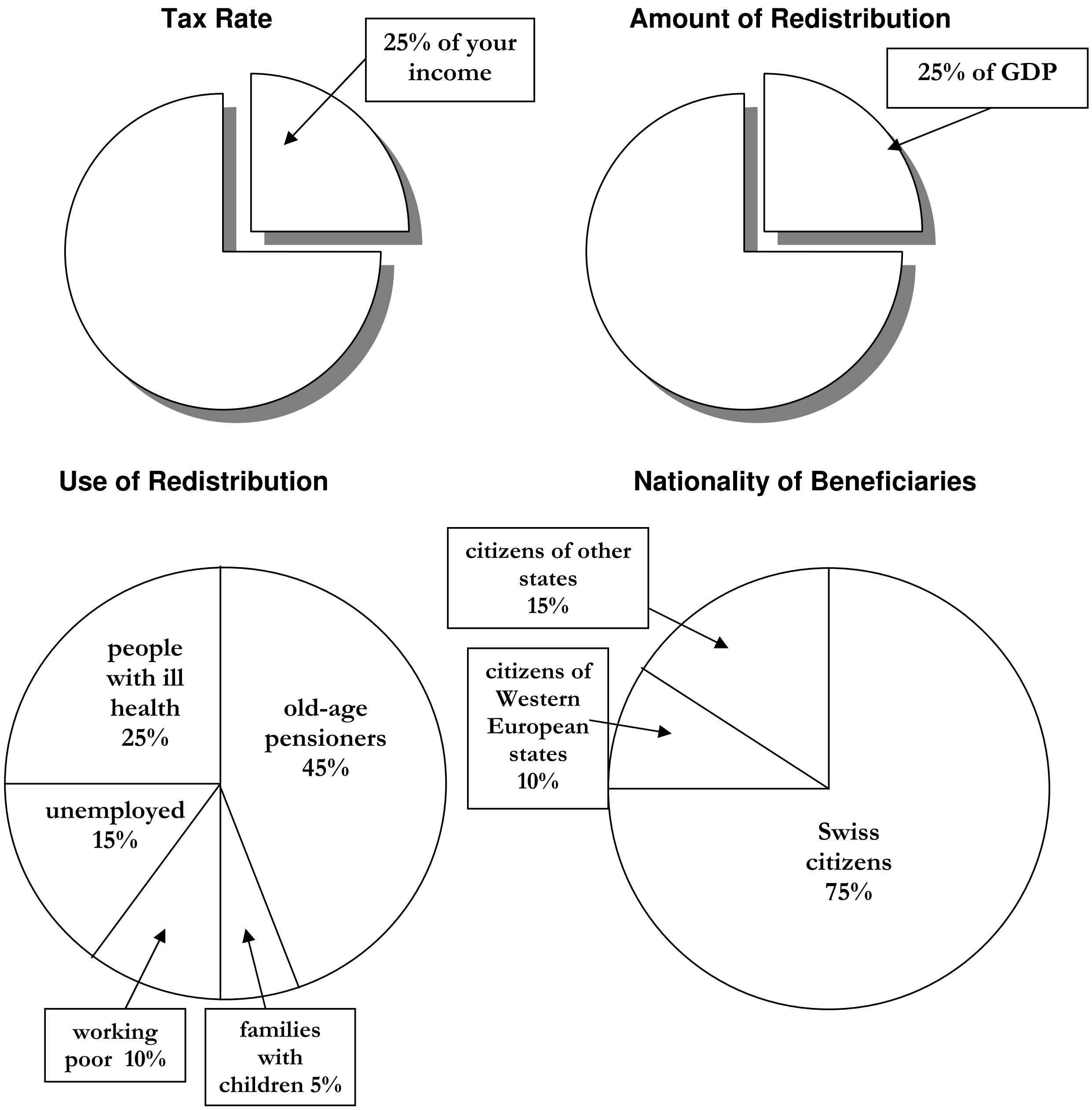


Exhibit 2: Card for Alternative No. 1

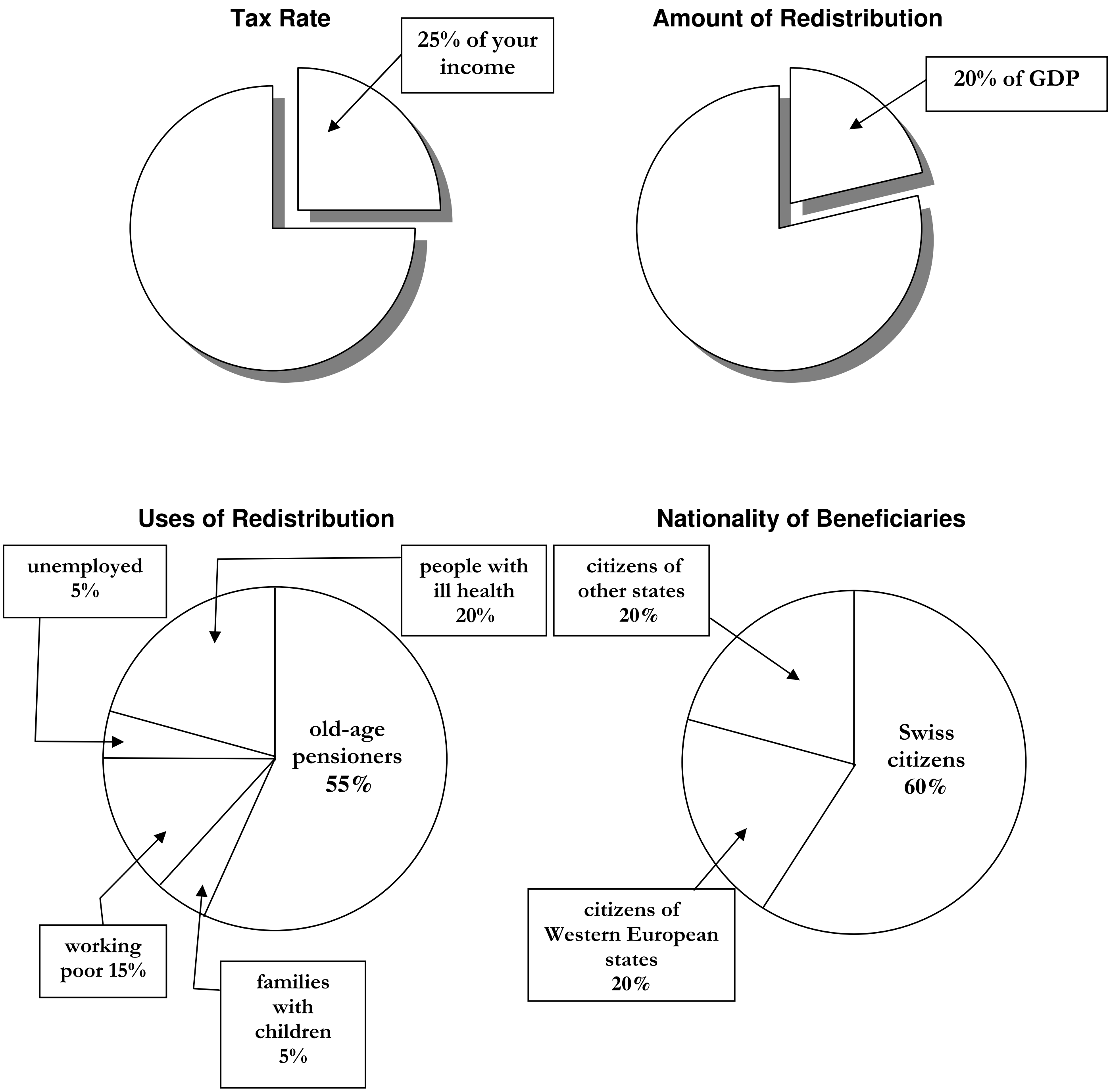
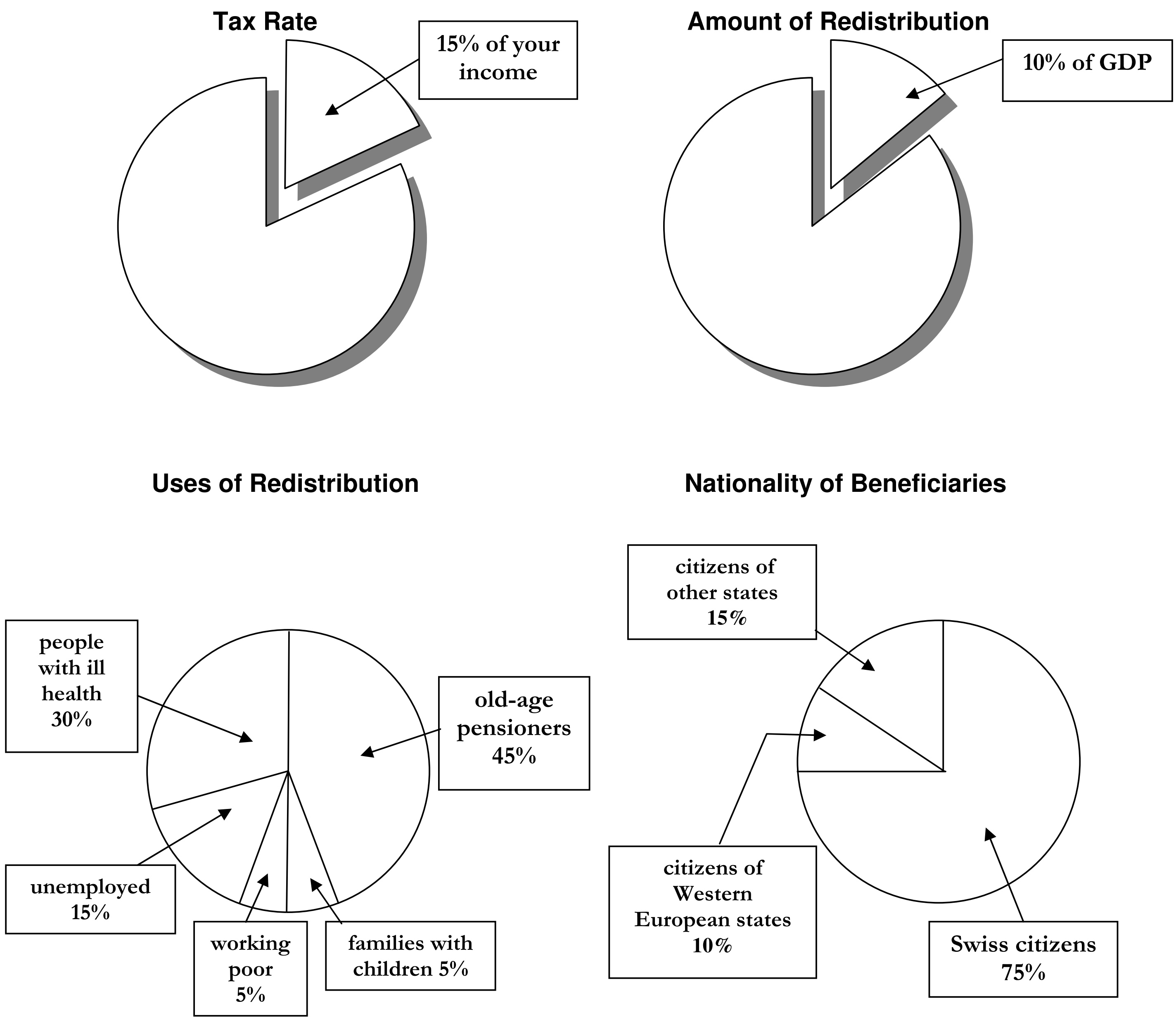


Exhibit 3: Card for Alternative No. 2



Is the welfare state sustainable? Experimental evidence on citizens' preferences for redistribution

Ilja Neustadt and Peter Zweifel

Abstract: The sustainability of the welfare state ultimately depends on citizens' preferences for income redistribution. They are elicited through a Discrete Choice Experiment performed in 2008 in Switzerland. Attributes are redistribution as GDP share, its uses (the unemployed, old-age pensioners, people with ill health etc.), and nationality of beneficiary. Estimated marginal willingness to pay (WTP) is positive among those who deem benefits too low, and negative otherwise. However, even those who state that government should reduce income inequality exhibit a negative WTP on average. The major finding is that estimated average WTP is maximum at 21% of GDP, clearly below the current value of 25%. Thus, the present Swiss welfare state does not appear sustainable.

Keywords: Income redistribution, welfare state, sustainability, preferences, willingness to pay, discrete choice experiments

JEL classification: C35, C93, D63, H29

Chapter 5

Is the welfare state sustainable? Experimental evidence on citizens' preferences for redistribution

5.1 Introduction

The sustainability of the welfare state is a hotly debated topic between politicians and interest groups. The economists' contribution to the debate traditionally has been to analyze the effects of redistributive policies on employment, output, and growth. However, in full cognition of these effects, a majority of citizens may still exhibit willingness to pay (WTP) for more redistribution of income. Conversely, its WTP may be negative even in a situation where these side effects of redistribution are unimportant. Ultimately, the sustainability of the welfare state therefore hinges on citizens' WTP. Through a Discrete Choice experiment (DCE), this paper seeks to determine not only the desired amount of redistribution but also to test several hypotheses concerning the determinants of this WTP. The data come from a DCE performed in the fall of 2008 and involving more than 900 Swiss citizens.

Recently, there has been a great deal of research into the demand for redistribution and its determinants, which will be discussed in Section 5.2 below. One strand relates the measured amount of redistribution to economic, institutional, and behavioral factors. Examples are Alesina and Giuliano (2009) and Akkoyunlu et al. (2009). However, the observed amount of redistribution is the outcome of an interaction between demand and supply, with supply governed by a country's political institutions and processes. This classical identification problem would have to be addressed in order to make inferences about citizens' preferences for redistribution. A second strand of research, exemplified by Alesina and La Ferrara (2005) and Guillaud (2008), relies on surveys designed to measure attitudes towards redistribution. The problem with this approach is its failure to impose a budget constraint. It therefore cannot predict actual decision making (e.g. voting at the polls), where citizens take the consequences in terms of their own income and wealth into account. A third approach seeks to solve this problem through Contingent Valuation (CV) experiments [see e.g. Boeri et al. (2001, 2002)[1]]. The weakness of the CV approach is that it holds all the attributes of the good in question constant, varying its price only. One would want to vary other attributes of redistribution besides its tax price, viz. its uses (for health, old age, etc.) and the type of beneficiary (foreigner, national).

By way of contrast, a DCE allows to measure preferences uncontaminated by supply influences, it imposes the budget constraint through the price attribute, and it does so in a realistic way by making respondents choose between alternatives where all attributes are allowed to vary.

The remainder of this paper is structured as follows. Section 5.2 contains a literature review from which hypotheses to be tested are derived. Its first part concerns general determinants of the demand for redistribution and the second, attitudes towards reduction of inequality as determinants of preferences for redistribution. Section 5.3 presents a general description of the method of DCEs as well as the design of the present experiment. The descriptive statistics of the experiment follow in Section 5.4, and hypothesis tests, in Section 5.5. Section 5.6 summarizes the results and concludes with an assessment of the sustainability of the Swiss welfare state.

[1]Boeri et al. (2001) study international attitudes towards redistribution with a focus on pension and unemployment schemes in France, Germany, Italy, and Spain. They also perform CV experiments that impose an explicit trade-off between income and social insurance coverage on respondents. They find that people oppose an extension of the welfare state, with conflicts between young and old, rich and poor, and insiders and outsiders creating significant hurdles to welfare reform.

5.2 Literature review and statement of hypotheses

This section first presents research that defines the general background of this paper and then moves on to contributions that lead to a set of specific hypotheses to be tested.

5.2.1 General determinants of the demand for income redistribution

In their reviews, Alesina and Giuliano (2009) and Akkoyunlu et al. (2009) identify a wide set of factors that can be categorized as economic, political, and behavioral determinants of the demand for income redistribution.

Economic determinants

The simplest framework for the analysis of purely economic determinants is provided by a model focusing on current economic well-being, originally proposed by Romer (1975) and Roberts (1977) and extended by Meltzer and Richard (1981) [RRMR model]. This model assumes non-altruistic utility-maximizing individuals differentiated by their income levels only. The government pays a lump-sum transfer to all citizens, financed by a linear uniform income tax. Individuals with an income below the mean favor taxation and transfers while those with an income above the mean oppose it. In a political equilibrium, the majority of voters supports a positive tax rate corresponding to the value desired by the median voter[2]. The model's prediction is that the larger the gap between the mean and the median income, the higher the level of taxation and redistribution.

The empirical evidence is quite mixed. On the one hand, Alesina and Rodrik (1994), Persson and Tabellini (1994), and Milanovic (2000) find some supporting evidence. Furthermore, Guillaud (2008), conducting a cross-section analysis of survey data from four EU countries, shows that poorer and less educated individuals are more in favor of redistribution. On the other hand, Alesina and Glaeser (2004), Perotti (1996), and Rodriguez (1999) fail to find supporting evidence for this model. Moreover, Neustadt and Zweifel (2009) [see Chapter 3] relate willingness to pay (WTP) for income redistribution elicited from a Discrete Choice Experiment (DCE, see Section 5.3.1 for details) to measures of economic well-being. WTP

[2] The median voter's income is assumed to be below the mean. This assumption is satisfied for most economies.

values are shown to be negatively related to income and education, contradicting the RRMR model.

Another economic explanation is the "Prospect of Upward Mobility" (POUM) hypothesis, suggested by Hirschman and Rothschild (1973) as the 'tunnel effect' and more recently reformulated by Benabou and Ok (2001). It extends the RRMR model by introducing individuals' expectations, based on their observations regarding the income mobility of others in society. Expected upward mobility may dampen a poor but forward-looking voter's enthusiasm for income redistribution.

Empirical support of the POUM hypothesis is provided by Alesina and La Ferrara (2005) who, using an actual mobility matrix for the United States, show that people who can expect high future income oppose redistribution[3]. Rainer and Siedler (2008) use probabilistic expectations data to show that individuals with a sufficiently large chance of occupational upward mobility exhibit a lower demand for redistribution; conversely, those with a sufficiently large risk of occupational downward mobility opt for more redistribution. Checchi and Filippin (2004), testing the POUM hypothesis by means of a within-subjects experiment, find corroborating evidence under several alternative specifications. According to Guillaud (2008), however, individuals who subjectively experienced upward mobility over ten years tend to be more (rather than less) supportive of redistributive policies. Moreover, upward intergenerational mobility in occupational prestige goes along with more positive rather than negative attitude towards redistribution. Alesina and Giuliano (2009) examine the empirical evidence for the United States and briefly across countries, concluding that social mobility (if measured as the change in the occupational prestige) does decrease demand for redistribution once sociodemographic (age, gender, race) and socioeconomic characteristics (income, education) are controlled for. In their DCE-based study, Neustadt and Zweifel (2009) [Chapter 3] relate preferences for redistribution to mobility. They find partial empirical support for the POUM hypothesis.

Another economic explanation, suggested by the social contract literature, is that preferences for redistribution can at least in part be interpreted as a demand for insurance by

[3]The 'tunnel effect' also works in the opposite direction, causing forward-looking agents with high incomes but downward mobility expectations to be in favor of redistribution. This prediction is confirmed by Ravallion and Lokshin (2000) using a data set from Russia. Furthermore, Molnár and Kapitány (2006a,b) show that individuals who lack clear expectations about their future income favor redistribution even more than those with negative but clear expectations.

risk-averse individuals. In a hypothetical situation, where individuals do not yet know their endowment as well as their future position in society ['veil of ignorance', cf. Rawls (1999)], they are predicted to exhibit positive WTP for an income transfer from more favorable future states to less favorable ones. Redistributive policies can thus be interpreted as reflecting this hypothetical demand for insurance. Beck (1994) investigates individual behavior under the 'veil of ignorance' in an experiment. By placing participants in a hypothetical society with random differences in income, represented by lotteries, he is able to derive the desired amount of income redistribution. Individuals indeed display risk aversion, albeit not of the extreme kind implied by the Rawlsian maximin rule. Furthermore, they show no preference for income redistribution in excess of what can be explained by risk aversion.

Political Determinants

As to the political determinants of the demand for income redistribution, the literature [Persson and Tabellini (2000, 2003), Lizzeri and Persico (2001), Milesi-Ferretti et al. (2002)] predicts that proportional representation tends towards universal programs benefitting various groups (old-age pensioners, working poor, minorities, etc.), while majority rule results in targeted 'pork barrel' programs. Persson and Tabellini (2003) find supporting empirical evidence in that countries with proportional representation have GDP shares of government expenditure that ceteris paribus are 5 percentage points higher than countries with majority rule. Moreover, Akkoyunlu et al. (2009) [Chapter 2] present weak evidence of a positive correlation between the degree of proportional representation and the transfer share in GDP in OECD countries. Additional political determinants of redistribution include two-party vs. multiparty system, presidential vs. parliamentary democracy, and direct vs. representative democracy, with two-party systems, presidential, and direct democracies all predicted to induce less public redistribution. In order to sketch the institutional background of the DCE described in Section 5.3.2, Switzerland can be described as follows. It has a high degree of proportional representation and a parliamentary democracy. Its distinguishing feature, however, is its extensive direct democratic control in the guise of popular initiatives and referenda. This might serve to limit public welfare spending while enforcing efficiency in redistribution [cf. Feld et al. (2007)].

Behavioral Determinants

The mixed empirical evidence bearing on the economic determinants of preferences for redistribution calls for a detailed analysis of their behavioral determinants. In particular, beliefs have been at the center of attention. The theoretical base is laid by Alesina and Angeletos (2005), who develop a model where society's belief whether effort or luck determines economic success gives rise to multiple self-fulfilling equilibria. Benabou and Tirole (2006) propose a model for the emergence and persistence of such collective beliefs. Moreover, beliefs can be seen as a source of altruistic preferences and inequality aversion [see Section 5.2.2]. On the empirical side, Fong (2001) presents evidence in line with Alesina and La Ferrara (2005) suggesting that beliefs about the role of luck in determining economic success are an important determinant of the demand for redistribution. She also considers the effects of incentives. If effort determines income, then an increased income tax rate causes an output loss due to its effect on incentives. This consideration is hypothesized to qualify the link between beliefs and the demand for redistribution. However, the data fail to support this hypothesis. Neustadt (2010) [see Chapter 4] analyzes the influence of behavioral determinants on willingness to pay (WTP) for income redistribution in Switzerland and shows that the less religious citizens exhibit a significantly higher marginal WTP than the more religious ones. Further, the citizens without religious affiliation have a positive marginal WTP while those belonging to some denomination, a negative one.

5.2.2 Attitudes towards reduction of inequality and demand for income redistribution

While the POUM hypothesis suggests less redistribution than predicted by the RRMR model, the assumption of altruistic preferences can lead to the opposite prediction. In fact, if individuals care also about the utility of others, one might expect more redistribution than predicted by the conventional RRMR model. Fehr and Schmidt (2006) provide a review of several models of social preferences, in particular, altruism, envy, inequality aversion, fairness, and reciprocity. Here, we focus on inequality aversion to derive hypotheses relating it to demand for income redistribution. In a simple model of inequality aversion, Fehr and Schmidt (1999) assume that individuals feel envy if their incomes are below that of others [disadvantageous inequality, see second term of eq. (5.1)], but they feel altruistic when their income exceeds it [advantageous inequality, see third term of eq. (5.1)]. An individual i's utility function is assumed to have the form

$$U_i(x_1,\ldots,x_N) \quad = \quad x_i - \frac{\alpha_i}{N-1}\sum_{j\neq i}\max\{x_j - x_i, 0\} - \frac{\beta_i}{N-1}\sum_{j\neq i}\max\{x_i - x_j, 0\} \tag{5.1}$$

with $0 \leq \beta_i \leq \alpha_i$ (the disutility from disadvantageous inequality is assumed to exceed that from advantageous inequality) and $\beta_i \leq 1$ (individuals are not willing to waste money in order to avoid being significantly richer than others). Here x_k, $k = 1,\ldots,N$, denotes individual k's income, α_i, the marginal disutility from disadvantageous inequality, and β_i, the marginal disutility from advantageous inequality. In this model, the decisive median voter demands more redistribution than in the conventional RRMR model. First, she has disutility from being richer than those with income $x_j < x_i$. Second, she has even more disutility from being poorer than those with income $x_j > x_i$. Thus, in a political equilibrium, larger values of α_i, β_i (higher level of inequality aversion) lead to a higher demand for redistribution compared to that predicted by the RRMR model.

Based on the assumption of inequality aversion, we formulate two hypotheses to be tested in Section 5.5.2. The first assumes that citizens with higher inequality aversion tend to deem the current level of social benefits to be too low, while those with lower inequality aversion deem it to be too high or just sufficient. Thus, the former are predicted to exhibit a positive WTP for redistribution while the latter, a negative one. The second hypothesis is based on the consideration that voters exhibiting inequality aversion tend to support the view that the government should reduce the income gap between the poor and the rich. Consequently, respondents who state that the reduction of the income gap is a task of the government are expected to exhibit a positive WTP for redistribution.

HYPOTHESIS 1: *Willingness to pay for redistribution is expected to be*

(A) ***negative*** *if the currently provided level of social benefits is considered* ***too high****,*

(B) ***negative but less so than in*** (A) *if the currently provided level of social benefits is considered to be* ***just sufficient****,*

(C) ***positive*** *if the currently provided level of social benefits is considered* ***too low****.*

Hypothesis 2: *Willingness to pay for redistribution is expected to be*

(a) ***negative*** *if the individual thinks that the government* **should not** *reduce the income gap between the poor and the rich,*

(b) ***positive*** *if the individual thinks that the government* **should** *reduce the income gap between the poor and the rich.*

5.3 Discrete choice experiments

5.3.1 Theoretical foundations

Discrete Choice Experiments (DCEs) provide a tool for measuring individuals' preferences for characteristics of commodities, the so-called attributes. In contradistinction to classical Revealed Preference Theory, originating with Samuelson (1938), DCEs allow individuals to express their preferences for non-marketed as well as hypothetical products. During a DCE, respondents are repeatedly asked to compare the status quo with several hypothetical alternatives defined by their attributes including a price. By varying the levels of attributes, different product alternatives are generated. A rational individual will always choose the alternative with the highest utility. From the observed choices, the researcher can infer the utility associated with the attributes. The proposed method, derived from the New Demand Theory of Lancaster (1971), is also known as Conjoint Analysis [Louviere et al. (2000)].
The most prominent alternative to a DCE is Contingent Valuation (CV). A certain situation or product is described in detail, and respondents are asked to indicate their maximum willingness to pay (WTP) for this fixed product. Only its price attribute is varied, while in Conjoint Analysis all relevant attributes are varied simultaneously, making it a multi-attribute valuation method [Merino-Castello (2003)]. While a DCE describes the product in less detail than a typical CV study, it allows for analyzing many product varieties by varying the levels of relevant attributes [Louviere et al. (2000), p. 344]. Trade-offs among attributes can be explicitly taken into account and WTP values of attributes estimated separately (see below). Furthermore, strategic behavior of respondents is less likely than in CV with its exclusive emphasis on price, which facilitates strategic behavior. Finally, biases that easily occur when individuals are directly asked about their WTP are less frequently observed in a DCE [Ryan (2004)].

A particular advantage of a DCE in the present context is that it permits to explicitly impose the budget constraint through a price attribute in the guise of the tax share of income used to finance the transfers considered. Respondents can be made to simultaneously choose this share and hence the 'size of the pie' and the 'slices of the pie' devoted to different types of recipients and uses (health, old age, etc.; see Exhibits No. 1 to 3 in Appendix). Thus, trade-offs among different attributes of the redistribution plan can be calculated to assess the relative importance of the respective redistributive goals.

The econometric method used is based on the Random Utility Theory [see Luce (1959), Manski and Lerman (1977), McFadden (2001, 1974, 1981, 2001)]. Individual i values alternative j according to the utility V_{ij} attained, which is given by

$$V_{ij} = v_i(a_j, p_j, y_i, s_i, \varepsilon_{ij}). \tag{5.2}$$

Here, $v_i(\cdot)$ denotes i's indirect utility function, a_j, the amount of attributes associated with alternative j, and p_j, price. The individual's income and sociodemographic characteristics are symbolized by y_i and s_i, respectively. Finally, ε_{ij} denotes the error term, which is due to the fact that the experimenter will never observe all the arguments entering v_i, imparting a stochastic element to observed choices. As usual, the utility function is additively split into a systematic component $w(\cdot)$ and a stochastic one,

$$V_{ij} = w_i(a_j, p_j, y_i, s_i) + \varepsilon_{ij}.$$

A utility-maximizing individual i will prefer alternative j to alternative l if and only if

$$w_i(a_l, p_l, y_i, s_i) + \varepsilon_{il} \leq w_i(a_j, p_j, y_i, s_i) + \varepsilon_{ij}. \tag{5.3}$$

Due to the presence of the stochastic term, only the probability P_{ij} of individual i choosing alternative j rather than alternative l can be estimated, with

$$P_{ij} = \text{Prob}\left[w_i(a_l, p_l, y_i, s_i) + \varepsilon_{il} \leq w_i(a_j, p_j, y_i, s_i) + \varepsilon_{ij}\right] \tag{5.4}$$

$$= \text{Prob}\left[\varepsilon_{il} - \varepsilon_{ij} \leq w_i(a_j, p_j, y_i, s_i) - w_i(a_l, p_l, y_i, s_i)\right]. \tag{5.5}$$

Thus, the probability of choosing j amounts to the probability of the systematic utility difference $w_i[j]-w_i[l]$ dominating the 'noise', $\varepsilon_{il}-\varepsilon_{ij}$. The error terms $\{\varepsilon_{il}, \varepsilon_{ij}\}$ can be assumed to be normally distributed with mean zero and variances σ_l^2 and σ_j^2 as well as covariance σ_{lj}. Under these assumptions, $\varphi_{ij} := \varepsilon_{il} - \varepsilon_{ij}$ is also normally distributed with mean zero and variance $\sigma^2 := \text{Var}[\varphi_{ij}] = \sigma_l^2 + \sigma_j^2 - 2\sigma_{lj}$. Thus, equation (5.5) can be represented as

$$P_{ij} \quad = \quad \Phi\left(\frac{w_i(a_j, p_j, y_i, s_i) - w_i(a_l, p_l, y_i, s_i)}{\sigma}\right), \tag{5.6}$$

where $\Phi(\cdot)$ denotes the cdf of a standard normal distribution. This model is known as the binary probit model [cf. Ben-Akiva and Lerman (1985)]. Hensher et al. (1999) provide empirical evidence that a linear specification of the function $w(\cdot)$ leads to good predictions in its middle ranges. Therefore, in the case of the simple model that relates utilities and choice probabilities to the attributes only (see Section 5.5.1), one posits

$$w_i(a_j, p_j, y_i, s_i) \quad = \quad c_i + \sum_{k=1}^{K} \beta_k a_{kj} + \varepsilon_{ij}, \tag{5.7}$$

where c_i represents an individual-specific constant, a_k, $k = 1, \ldots, K$, are the attributes of the alternative, and β_k, $k = 1, \ldots, K$, are the parameters to be estimated. These parameters can be interpreted as the constant marginal utilities of the corresponding attributes. One obtains the following expression representing the difference in utility of individual i between alternative j and status quo,

$$\Delta V_{ij} \quad = \quad c_i + \sum_{k=1}^{K} \beta_k \Delta a_{kj} + \beta_p \Delta p_j + \varphi_{ij}, \tag{5.8}$$

where $\Delta a_{kj} = a_{kj} - a_{lj}$, $\Delta p_j = p_j - p_l$, $c_i = c_{il} - c_{ij}$, and $\varphi_{ij} = \varepsilon_{il} - \varepsilon_{ij}$ for each $j \neq l$. The marginal rate of substitution between two attributes m and n is given by

$$\text{MRS}_{m,n} \quad = \quad -\frac{\partial v / \partial a_m}{\partial v / \partial a_n}. \tag{5.9}$$

Therefore, the marginal WTP for attribute a_m can be calculated by dividing the marginal utility of this attribute by the marginal utility of the price attribute [in the present context, the income tax rate, see e.g. Telser (2002), p. 56]:

$$\text{MWTP}(a_m) = \frac{\partial v/\partial a_m}{\partial v/\partial p_j}. \tag{5.10}$$

For econometric inference, it is important to recall that the same individual makes several choices. The two-way random-effect specification takes this into account with $\varphi_{ij} = \mu_i + \eta_{ij}$, where μ_i denotes the component that varies only across individuals but not across the choice alternatives. The terms μ_i and η_{ij} are assumed uncorrelated with the product attributes $(a_{i1}, \ldots, a_{iK})$ and between themselves. By a standard assumption in a probit model, $\sigma_\eta = 1$. Hence $\text{Var}[\varphi_{ij}] = \sigma_\eta^2 + \sigma_\mu^2 = 1 + \sigma_\mu^2$ and $\text{Corr}[\varphi_{ij}, \varphi_{il}] = \frac{\sigma_\mu^2}{1+\sigma_\mu^2} =: \rho$. The parameter ρ indicates how strongly the various responses of an individual are correlated with each other, or, equivalently, the share of the total variance that can be explained by individual-specific error term. The random-effects specification is justified if ρ is high and significant.

The simple model can be extended by including various socioeconomic variables (e.g. income group, level of education, social mobility). These variables need to be interacted with the product attributes as well as with the constant, giving rise to the extended model specification which allows to check for preference heterogeneity and thus to test Hypotheses 1 and 2 in Section 5.5.2. By means of a t test we can investigate whether the differences in marginal WTP values between different socioeconomic groups are statistically significant. The computation of the variance of the marginal WTP values is performed by the delta method, cf. Hole (2007).

5.3.2 Experimental design

A representative telephone survey with 979 respondents was conducted in the fall of 2008. Prior to the survey, the attributes and their levels used to define 'income redistribution' had been checked in two pretests for their relevance. Attributes form four groups (see Table 5.1).

1. Shares of the total redistribution budget to be spent on five types of recipients (viz. the working poor, the unemployed, old-age pensioners, families with children, and people with ill health);

Attribute	Label	Status Quo Level	Alternative Levels
Shares of benefits going to			
• Working Poor	W_POOR	10%	5%, 15%
• Unemployed	UNEMP	15%	5%, 25%
• Old-Age Pensioners	PENS	45%	35%, 55%
• Families with Children	FAM	5%	10%
• People with Ill Health	ILL	25%	20%, 30%
Shares of benefits going to			
• Swiss citizens	SWISS	75%	60%, 85%
• Western European foreigners	WEU_FOR	10%	5%, 20%
• Other foreigners	OTH_FOR	15%	10%, 20%
Total amount of redistribution	REDIST	25% (of GDP)	10%, 20%, 30%, 40%, 50%
Income tax	TAX	25% (of personal income)	10%, 15%, 40%

Table 5.1: Attributes and their levels

2. Shares of the total redistribution budget to be spent on three groups (viz. Swiss citizens, western European foreigners, and other foreigners);

3. Total amount of redistribution, defined as a share of GDP;

4. Personal income tax rate to be paid by the respondent (the price attribute).

Clearly, these attributes and their levels combine to form a total number of possible scenarios that cannot be realized in an experiment. The scenarios define the n rows of the observation matrix X, with associated covariance matrix $\Omega = \sigma^2 \left(X'X\right)^{-1}$ of parameters β to be estimated. So-called D-efficient design calls for the minimization of the geometric mean of the eigenvalues of Ω,

$$D \, \text{efficiency} = \left(|\Omega|^{\frac{1}{K}}\right)^{-1}$$

where K denotes the number of parameters to be estimated [cf. Carlsson and Martinsson (2003)]. Using this optimization procedure and incorporating several restrictions, the number of alternatives was reduced to 35 and randomly split into five groups. One alternative was included twice in each decision set for a consistency test, resulting in 8 binary choices per respondent.

In order to make sure that decisions were based on a homogeneous information set and made in a consistent way, respondents were provided with a detailed description of the attributes and their possible realizations. The Appendix shows the graphical representation of the status quo (Exhibit 1) and two selected alternatives (Exhibits 2 and 3).

5.4 Descriptive statistics

5.4.1 Socioeconomic characteristics

	too little		right amount		too much		total valid answers		missing
Income group, CHF[a]	No.	%	No.	%	No.	%	No.	%	No.
< CHF 2000	63	35	100	56	16	9	179	100	13
CHF 2000 - 3999	58	32	94	53	27	15	179	100	14
CHF 4000 - 5999	141	43	149	45	39	12	329	100	15
≥CHF 6000	79	37	118	56	14	7	211	100	10
Missing	11		16		1		28		
Total answers	352	38	477	52	97	10	926		53

[a]1 CHF (Swiss franc) = 0.8 US$ at 2008 exchange rates

Table 5.2: Answers to the question "Do you think that the government is spending too much, too little or about the right amount on welfare?", by income group

	yes		no		total valid answers		missing
Income group, CHF[a]	No.	%	No.	%	No.	%	No.
< CHF 2000	78	42	108	58	186	100	6
CHF 2000 - 3999	112	59	77	41	189	100	4
CHF 4000 - 5999	124	37	212	63	336	100	8
≥CHF 6000	90	42	122	58	212	100	9
Missing	13		16		29		
Total answers	417	45	535	55	952		27

[a]1 CHF (Swiss franc) = 0.8 US$ at 2008 exchange rates

Table 5.3: Answers to the question "Do you agree with the following statement: *'By increasing the income tax rates for rich families and financially supporting poor families, the government should try to reduce the income gap between the rich and the poor'*?", by income group

The sample consists of 979 respondents, 70 percent of them residing in the German-speaking part and 30 percent in the French-speaking part of Switzerland. Some 94 percent are born in the country, 50 percent are men, 20 percent having a monthly income below CHF 2,000 and 23 percent, above CHF 6,000, reflecting the structure of the Swiss population. However, only 1.5 percent of the respondents are unemployed.

38 percent of the respondents stated that the current level of social benefits was too low, 10 percent stated that it was too high, and 52 percent found it exactly right [see Table 5.2]. On the other hand, 45 percent of the respondents agreed with the statement, 'By increasing the income tax rates for rich families and financially supporting poor families, the government

	yes		no		total valid answers		missing
Income classes, CHF[a]	No.	%	No.	%	No.	%	No.
insurance	164	33	339	67	503	100	10
inequality reduction	219	55	181	45	400	100	13
Missing	34		15		49		
Total answers	417	45	535	55	952		27

[a]1 CHF (Swiss franc) = 0.8 US$ at 2008 exchange rates

Table 5.4: Answers to the questions "Do you agree with the following statement: *'By increasing the income tax rates for rich families and financially supporting poor families, the government should try to reduce the income gap between the rich and the poor'*?" and "What is your main motive for redistribution: insurance or inequality reduction?"

should try to reduce the income gap between the rich and the poor', while 55 percent disagreed [see Table 5.3].

The distribution of answers over income groups of the respondents is obviously in contradiction with the RRMR model. For instance, 35% of respondents with monthly incomes below CHF 2,000 (the 'poor') deem the current amount of social benefits too low, but this holds true for even 37% of those with incomes above CHF 6,000 (the 'rich') [see Table 5.2]. Similarly, the percentage of those finding the current size of the welfare state excessive is 9% among the 'poor' but only 7% among the 'rich'. Moreover, the share of those supporting a reduction of the income gap by public redistribution is 42% both among the 'rich' and the 'poor' [see Table 5.3]. Obviously, beliefs do not correlate with income. On the other hand, they may reflect inequality aversion. These findings motivate examining explanations of the demand for income redistribution based on beliefs and inequality aversion. However, as noted in Section 5.2.1, inequality aversion could be due to risk aversion in front of the 'veil of ignorance'. Indeed, 56 percent of the respondents state 'insurance' as their main motive for redistribution, compared to 44 percent of those with the 'inequality reduction' motive [see Table 5.4]. Attitudes clearly differ between the two groups, too. Only one-third of respondents with the 'insurance' motivation support the idea of inequality reduction to be provided by government, compared to 55% of those with the 'inequality reduction' motivation. In sum, 'true' inequality aversion in the sense of Fehr and Schmidt (1999) may well be relevant, at least in the present sample.

5.4.2 Respondents' choice behavior

There is a total of $979 \cdot 8 = 7,832$ decisions, of which not quite 20 percent were made in favor of an alternative over the status quo [see Table 5.5]. There are at least three explanations for this low percentage. First, in spite of checking in the pretests, the levels of the attributes in the experiment may not have been sufficiently spaced apart to make respondents switch. Second, some attributes (e.g. benefits going to the unemployed; see Table 5.7), may not have been important enough to cause a switch. Finally, there may be errors in decision making because the consistency test revealed 14 percent of choices to be inconsistent. However, there may simply be marked status quo bias in the face of highly complex decision-making situations, as suggested by the large negative constant in Table 5.7. Nonetheless, only 21 percent of respondents never opted for an alternative [see Table 5.6]. Conversely, almost 80 percent departed from the status quo at least once.

Choices	No.	in percent
for alternative	1,562	19.94
for status quo	6,088	77.73
No decision	182	2.32
Total	7,832	100

Table 5.5: Total number of choices

# choices for alternative	No.	in percent
0	209	21.35
1	309	31.56
2	226	23.08
3	131	13.38
4	57	5.82
5	16	1.63
6	10	1.02
7	0	0.00
8	5	0.51
Total valid answers	965	98.57
Missing	14	1.43
Sample	979	100

Table 5.6: Distribution of the number of chosen alternatives per respondent

Variable	Coeff.	Std. err.	z	$P > \|z\|$	Marg. eff.
Recipients' Social Group					
W_POOR	0.02784	0.00714	3.90	0.000	0.00697
UNEMP	0.01134	0.00452	2.51	0.012	0.00284
ILL	0.01600	0.00463	3.46	0.001	0.00400
FAM	0.06378	0.00942	6.77	0.000	0.01596
Recipient's Nationality					
SWISS	0.03656	0.00552	6.63	0.000	0.00915
WEU_FOR	0.02925	0.00869	3.37	0.001	0.00732
REDIST	-0.00523	0.00176	-2.97	0.003	-0.00131
REDIST^2	-0.06619	0.01174	-5.64	0.000	-0.01656
TAX	-0.02053	0.00183	-11.21	0.000	-0.00514
CONSTANT	-1.29878	0.06132	-21.18	0.000	n.a.

# observations	7,650
Log likelihood	-3,566.76
$\chi^2(0)$	108.87
Prob $> \chi^2$	0.000
σ_u	0.41610
ρ	0.14759

Table 5.7: Random effects probit estimates for the simple model

5.5 Estimation results

5.5.1 Simple model: preferences of the average respondent

Estimation of equation (5.8) includes REDIST^2 to allow for a possible nonlinearity of the indirect utility function with regard to the GDP share of redistribution REDIST. Moreover, the fact that uses and types of beneficiaries add up to 100 percent needs to be taken into account [see Table 5.1]. In order to avoid perfect collinearity, PENS (Pensioners) and OTH_FOR (Other foreigners) were dropped to obtain

$$\begin{aligned}\Delta V \quad = \quad & c_0 + \beta_1 \text{W\_POOR} + \beta_2 \text{UNEMP} + \beta_3 \text{ILL} + \beta_4 \text{FAM} + \\ & + \gamma_1 \text{SWISS} + \gamma_2 \text{WEU\_FOR} + \\ & + \delta_1 \text{REDIST} + \delta_2 \text{REDIST}^2 + \eta \text{TAX} + \varphi\end{aligned} \tag{5.11}$$

Estimation of a few of the $5 \cdot 3 = 15$ specifications with alternative exclusions produced results similar to those displayed in Table 5.7. Specifically, they agree in that alternatives with additional redistribution are chosen with a lower probability [for details with regard to ‘slices’ of the pie, see Neustadt and Zweifel (2010a)]. Also, note the sizeable and highly significant coefficient of the price attribute TAX, which is important for the estimation of marginal willingness-to-pay (MWTP) values [see eq. (5.10)]. For redistribution, the MWTP value is given by

$$\mathrm{MWTP}_{\mathrm{REDIST}} = \frac{\partial \Delta V / \partial \mathrm{REDIST}}{\partial \Delta V / \partial \mathrm{TAX}} = -\frac{\delta_1 + 2\delta_2 \mathrm{REDIST}}{\eta} \tag{5.12}$$

This amounts to -0.25 percentage points of income share per additional percentage point of GDP devoted to redistribution in excess of the status quo. Evaluated at the mean personal income of the sample, this equals CHF -11.78 per month. However, this figure is dwarfed by the compensation one would have to pay respondents to depart from the status quo, amounting to an estimated 63 percent of their monthly income, or 5.27 percent of their annual income [see the large negative constant in Table 5.7].

Equation (5.12) serves as the basis for checking the sustainability of the welfare state. Construction of the (quadratic) WTP function yields a maximum (with MWTP=0) at 21.05% of GDP, definitely below the current value of 25%. Therefore, the Swiss welfare state can be said to be too big in the light of average citizens’ preferences.

5.5.2 Extended model: preference heterogeneity

Ex-ante evaluation of the current level of social benefits and preferences for redistribution

The simple model is now extended by one attitudinal variable at a time. The first is respondents’ ex-ante evaluation of the current level of social benefits [SB, see Table 5.2]. The three levels of SB are represented by two dummy variables, SB_TOOHI and SB_TOOLOW. For instance, the latter is defined as

$$\mathrm{SB\_TOOLOW} = \begin{cases} 1 & \text{if the current level of benefits is deemed too low} \\ 0 & \text{otherwise.} \end{cases}$$

	exp. sign	MWTP, % of income	MWTP, CHF	s.e., CHF	
Social benefits too high (Group A)	-	-0.55946	-26.75	16.70	***
The right amount (Group B)	≈0	-0.41789	-19.61	8.34	***
Social benefits too low (Group C)	+	0.05487	2.47	8.09	

Note: *** denotes statistical significance of MWTP in % of income at the 1 percent level.

Table 5.8: Marginal WTP values for redistribution (in percent of monthly personal income and CHF) derived from the extended model with ex-ante evaluation of the current level of social benefits

The reference category is SB_RIGHT, indicating that the respondent deemed social benefits to have the right size. Since an attribute's marginal utility may vary with attitude, eq. (5.11) is modified to also contain interaction terms involving the attitudinal variables, resulting in

$$\begin{aligned}\Delta V' &= c_0' + \dots + \alpha_1' \mathrm{SB\_TOOLOW} + \dots + \alpha_2' \mathrm{REDIST} + + \alpha_3' \mathrm{REDIST}^2 + \dots \\ &\quad + \lambda_2' \mathrm{REDIST} \cdot \mathrm{SB\_TOOLOW} + \lambda_3' \mathrm{REDIST}^2 \cdot \mathrm{SB\_TOOLOW} + \dots \\ &\quad + \lambda_4' \mathrm{REDIST} \cdot \mathrm{SB\_TOOHI} + \lambda_5' \mathrm{REDIST}^2 \cdot \mathrm{SB\_TOOHI} + \varphi'.\end{aligned} \tag{5.13}$$

Hypothesis 1(A) states that the demand for redistribution is expected to be negative if the currently provided level of social benefits is considered too high. It is confirmed, with the MWTP for one percentage point increase of the total amount of redistribution being a negative CHF -26.75 [see Table 5.8]. Hypothesis 1(C), stating that the demand for redistribution should be positive if the level of social benefits is considered insufficient, finds some empirical support by a positive but insignificant MWTP value of CHF 2.47. However, Hypothesis 1(B), predicting the demand for redistribution to be negative but close to zero for individuals who deem the current level of benefits just sufficient, cannot be confirmed. In fact, the average respondent in this group exhibits a significantly negative MWTP for redistribution of CHF -19.61 per month. A t test indicates that the difference in MWTP values between respondent groups A and B is not significant, again contradicting Hypothesis 1(B).

As a check on the sustainability of the welfare state in the face of preference heterogeneity, group-specific WTP functions are constructed. Group A is found to have their maximum WTP at a GDP share of 15.89% devoted to redistribution. The values of Groups B and C are 18.45% and 25.52% of GDP, respectively. Therefore, attitudes with regard to the

	exp. sign	MWTP, % of income	MWTP, CHF	s.e., CHF	
should not reduce (Group a)	-	-0.34515	-16.68	6.35	***
should reduce (Group b)	+	-0.08417	-3.63	9.25	

Note: *** denotes statistical significance of MWTP in % of income at the 1 percent level.

Table 5.9: Marginal WTP values for redistribution (in percent of monthly personal income and CHF) derived from the extended model with the assessment whether the government should reduce the income gap between the rich and the poor

amount of social benefits do go along with heterogeneous preferences with regard to income redistribution. These discrepancies point to sharp conflicts of interest in the event that the amount of redistribution were to be reduced to the value preferred by the average citizen.

Assessment of the government's role in dealing with inequality and preferences for redistribution

Next, the simple model is extended by including the dummy variable GOV_REDUCE (=1 if the respondent thinks that the government should reduce the income gap between the rich and the poor, =0 otherwise) as well as its interactions with the attributes. Thus, eq. (5.11) is modified to read,

$$\begin{aligned}\Delta V'' &= c_0'' + \cdots + \alpha_1''\mathrm{GOV\_REDUCE} + \cdots + \alpha_2''\mathrm{REDIST} + \\ &+\alpha_3''\mathrm{REDIST}^2 + \cdots + \kappa_2''\mathrm{REDIST}\cdot\mathrm{GOV\_REDUCE} + \\ &+\kappa_3''\mathrm{REDIST}^2\cdot\mathrm{GOV\_REDUCE} + \cdots + \varphi''\end{aligned} \tag{5.14}$$

Hypothesis 2(a) states that the demand for redistribution is expected to be negative if a respondent believes that the government should not reduce the income gap between the rich and the poor. It is confirmed because MWTP in Group (a) is CHF -16.68 and statistically significant. Hypothesis 2(b) with its prediction for MWTP to be positive if a respondent wants the government to reduce the income gap cannot be confirmed. If at all, MWTP is negative in Group (b) (but lacks statistical significance).

Thus, individuals who stated support for inequality reduction by the government seem to exhibit inconsistent behavior by having a negative willingness to pay for this reduction. However, the framing of the question, "Do you agree with the following statement: *'By increasing* *the income tax rates for rich families and financially supporting poor families, the government should try to reduce the income gap between the rich and the poor'*?" did not evoke the trade-off between the reduction of the income gap and the respondent's own income. By way of contrast, the WTP values come from a Discrete Choice Experiment (DCE), where the budget restriction is inevitably present.

Addressing the sustainability issue once more, recall that the average respondent would prefer a share of GDP devoted to redistribution of 21% rather than the current value of 25%. However, construction of the group-specific WTP functions indicates that the optimal values of REDIST are again somewhat apart, with 19.21% of GDP for Group (a) and 24.09% for Group (b), respectively. Therefore, demand for income redistribution as measured by this DCE, while below the amount provided by the government, once more differs importantly between subpopulations, rendering a reform of the Swiss welfare state difficult.

5.6 Conclusion and discussion

In this paper, we elicited citizens' willingness to pay (WTP) for redistribution through a Discrete Choice experiment performed in 2008. Based on a simple model that relates choices to the attributes of redistribution only, the average Swiss citizen would have to be paid a compensation of CHF 11.78 (some US$ 9.40) per month (0.25 percent of monthly income) for an additional percentage point of GDP devoted to public redistribution. In addition, a very marked status quo bias would have to be overcome by payment of another 63 percent of monthly income.

Such an experiment also permits to test several hypotheses concerning the determinants of the demand for redistribution without any confounding supply-side influences. In particular, Hypothesis 1 states that it is negative (close to zero) among citizens who think that public welfare currently provided welfare is excessive (sufficient). An extended model that includes the pertinent attitudinal variable as a regressor yields confirming evidence for the 'excessive' component; however, the 'sufficient' component is also related to a negative WTP value, contradicting the hypothesis. Hypothesis 2 predicts that citizens who do (not) want government to reduce the income gap between the rich and the poor exhibit positive (negative) WTP for redistribution. Here, the extended version of the model supports the 'not' component of the hypothesis whereas those in favor of closing the gap fail to exhibit a positive WTP value. The major finding of the paper, however, is that estimated average WTP is maximum at 21% of GDP devoted to redistribution, clearly below the current value of 25%. Moreover, this value differs importantly depending on attitudes toward the desirable amount of redistribution and the government's role in dealing with inequality. Thus, there is reason for concern with regard to the sustainability of the Swiss welfare state.

The analysis presented in this paper is subject to several limitations. First, several behavioral explanations of the demand for redistribution (risk aversion, other beliefs, religiosity) were not tested. However, recent contributions to the field show that up to 90 percent of cross-country differences in public spending can be related to institutional and behavioral factors [see e.g. Alesina and Glaeser (2004), Akkoyunlu et al. (2009)]. Thus, future work should be devoted to finding out whether these factors also influence stated WTP for redistribution. A first step in this direction is done by Neustadt (2010) [see Chapter 4]. Second, the status quo bias found in this paper calls for more detailed analysis. To the extent that it reflects risk aversion, it should induce demand for redistribution - contrary to the results presented here. Finally, the evidence only relates to a point of time in one country and thus may be subject to transitory shocks and country-specific influences. Still, by appealing to citizens' stated preferences, the present contribution sheds some light on the question whether a welfare state laying claim to one quarter of the GDP is sustainable.

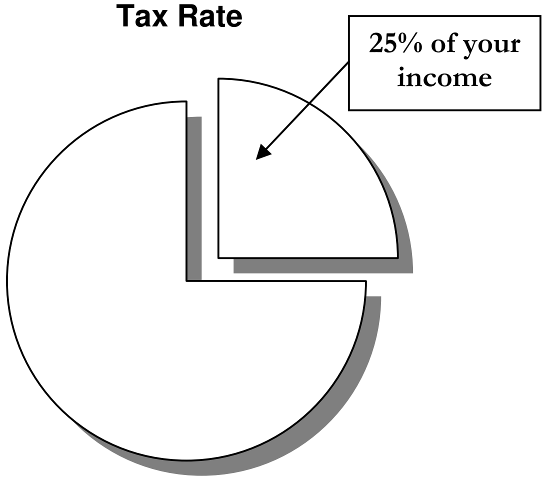
5.7 Acknowledgements

The authors gratefully acknowledge financial support from the Swiss National Science Foundation (SNF) under Project no. 100012-116398. They received helpful comments on earlier versions of this paper from Martin Beckmann, Douglas Bernheim, Joan Costa-i-Font, Mireia Jofre-Bonet, Mauro Gallegati, Reinhard Madlener, Erik Schokkaert as well as from participants in the conference 'From GDP to Well-Being: Economics on the Way to Sustainability' (Ancona, Italy, 3 to 5 December 2009) and the Workshop on Behavioural Welfare Economics (Venice Summer Institute, Venice, Italy, 21-22 July 2010), organized by Joan Costa-i-Font and Frank Cowell.

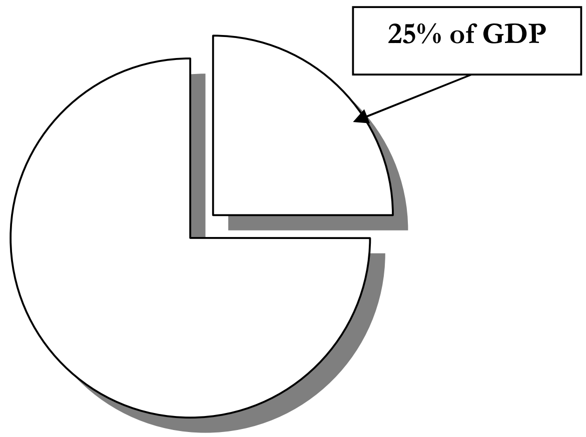
5.8 Appendix

Exhibit 1: Status Quo Card (current state of redistribution)

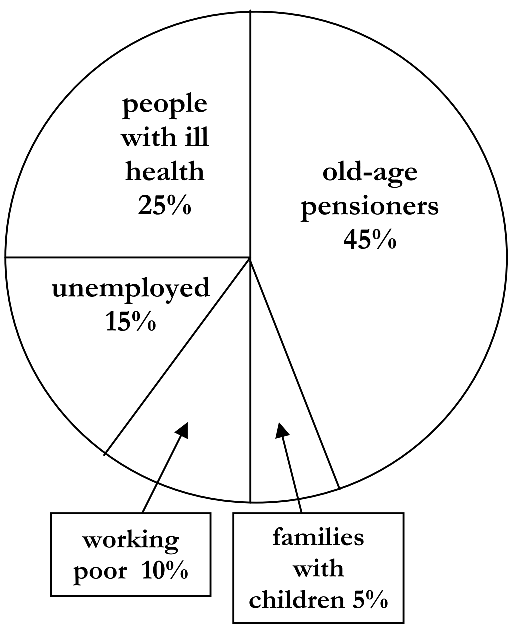
Tax Rate



Amount of Redistribution



Use of Redistribution



Nationality of Beneficiaries

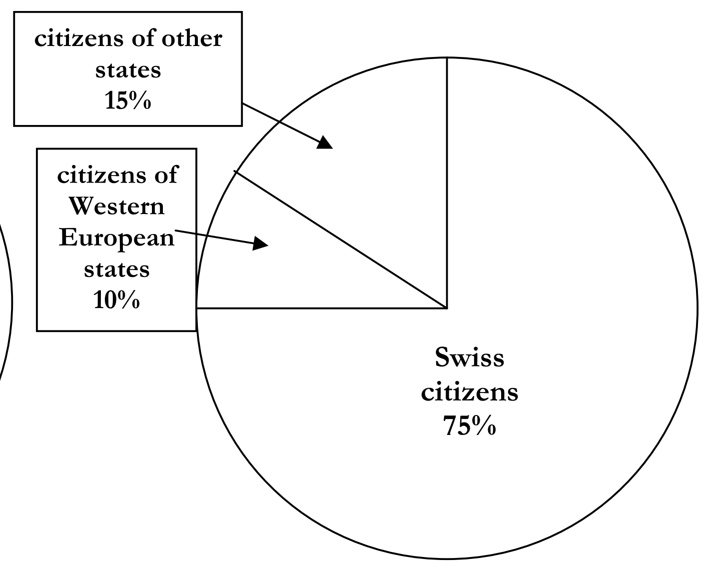


Exhibit 2: Card for Alternative No. 1

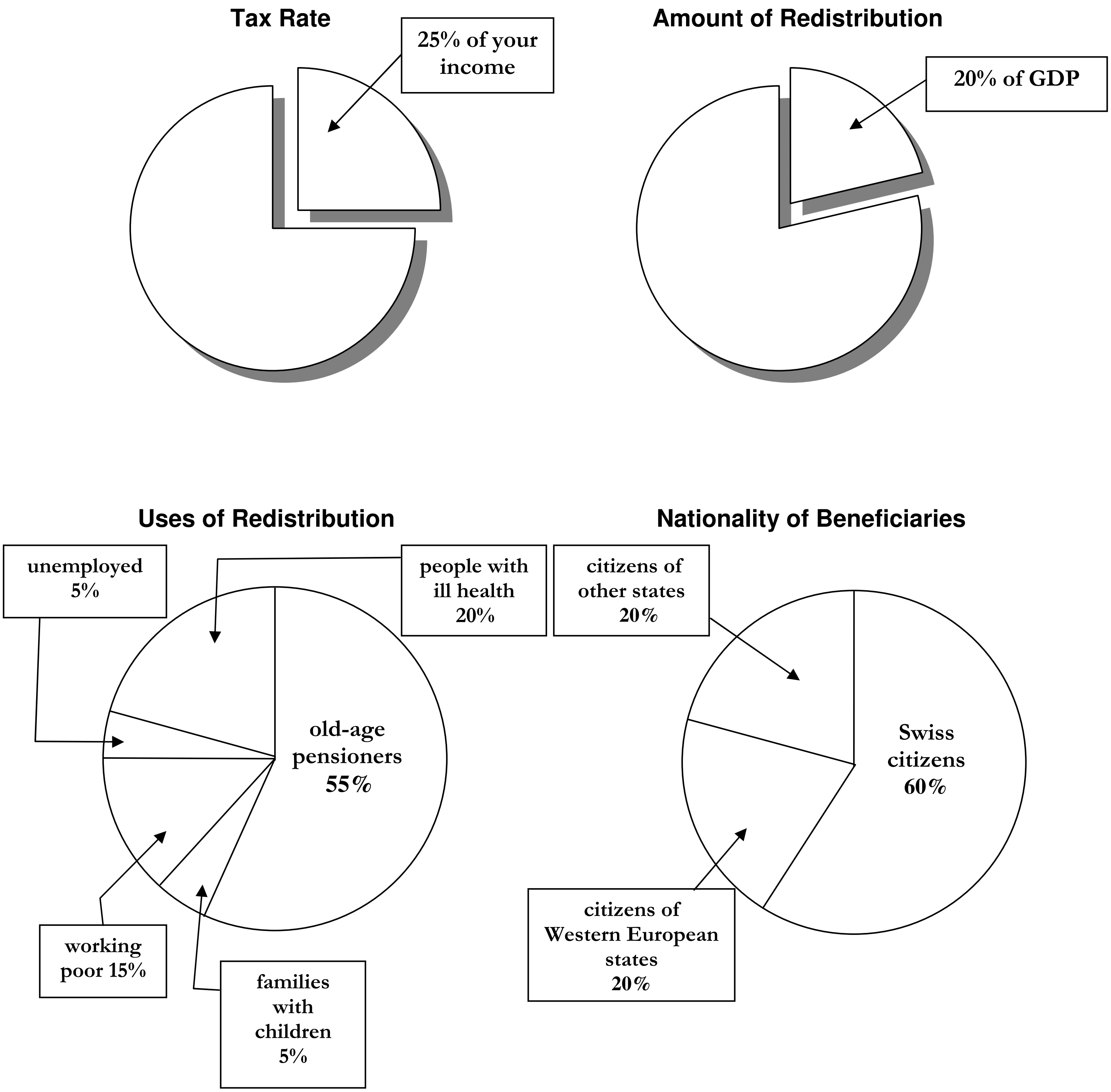
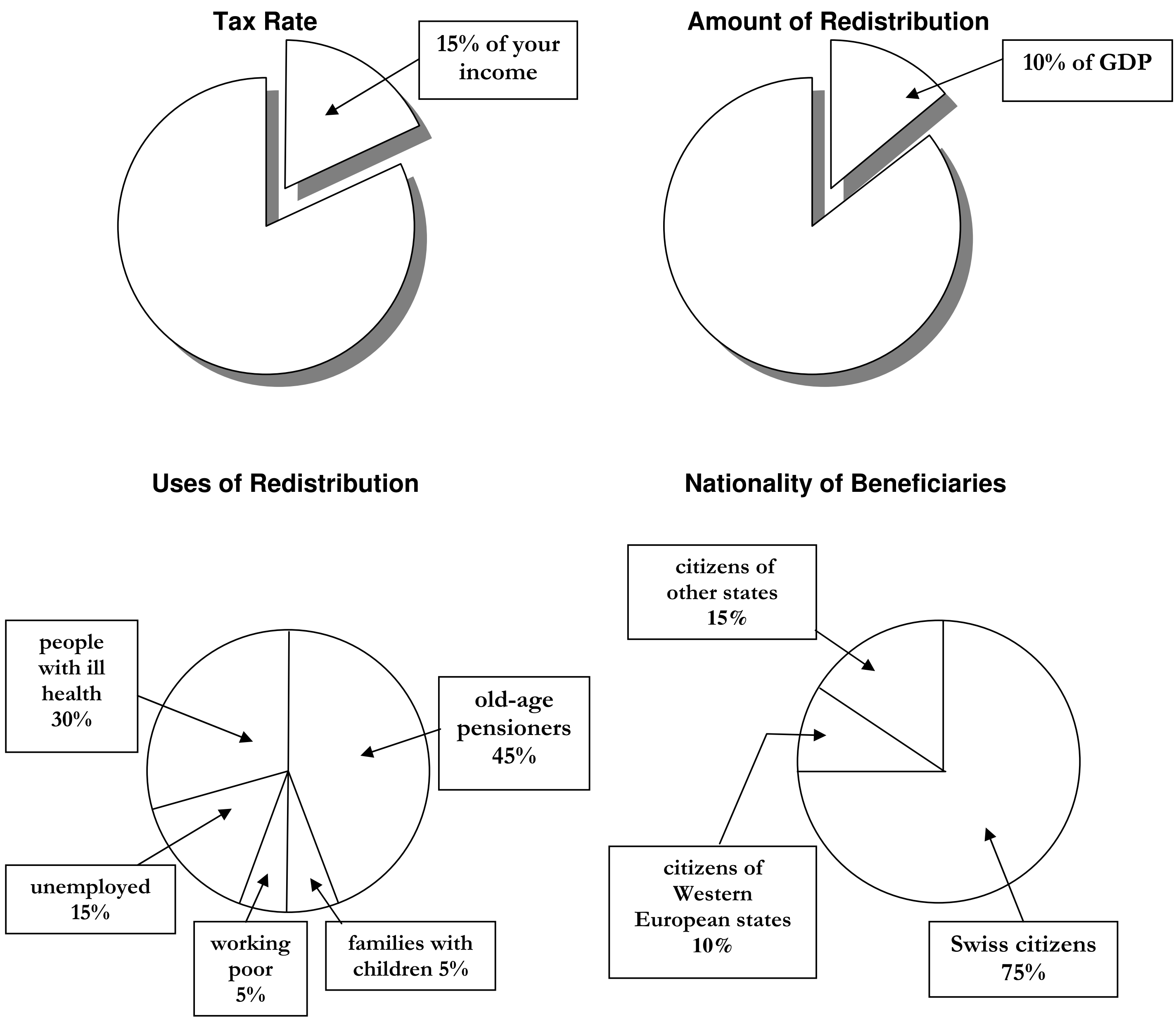


Exhibit 3: Card for Alternative No. 2



Promoting renewable electricity generation in imperfect markets: price vs. quantity control

Reinhard Madlener, Weiyu Gao,
Ilja Neustadt and Peter Zweifel

Revise and resubmit at "Energy Economics".

Abstract: The search for economically efficient policy instruments designed to promote the diffusion of renewable energy technologies in liberalized markets has led to the introduction of quota-based tradable 'green' certificate (TGC) schemes for renewable electricity. However, there is a debate about the pros and cons of TGC, a quantity control policy, compared to guaranteed feed-in tariffs, a price control policy. In this paper we contrast these two alternatives in terms of social welfare, taking into account that electricity markets are not perfectly competitive, and show that the price control policy dominates the quantity control policy in terms of social welfare.

Keywords: Green certificates; renewable portfolio standard; feed-in tariff

JEL classification: Q42, Q48

Chapter 6

Promoting renewable electricity generation in imperfect markets: price vs. quantity policies

6.1 Introduction

Electricity generation from renewable energy sources is increasingly recognized to play an important role for the achievement of a variety of primary and secondary energy policy goals, such as improved diversity and security of energy supply, reduction of local pollutant and global greenhouse gas emissions, regional and rural development, and exploitation of opportunities for fostering social cohesion, value added and employment at the local and regional level.

The plan of the European Commission of the 1990s to issue an EU Directive on the promotion of electricity from renewables [CEC (1998, 1999a,b)], which eventually led to the issuance of Directive 2001/77/EC [CEC (2001)], has triggered an intensive political and intellectual debate over the pros and cons of guaranteed feed-in tariffs (FIT) versus tradable green certificate (TGC) schemes [e.g. Rader (2000); Berry (2002); Lauber (2004); Palmer and Burtraw (2005); Madlener and Stagl (2005); Kildegaard (2008)].[1] Recently, a new and more comprehensive EU Draft Directive for renewable energy promotion has been published [CEC

[1]In the literature quota-based TGC schemes are sometimes also referred to as Renewable Portfolio Standards (RPS).

(2008a)], in which no clear preference for one or the other instrument is indicated. According to an accompanying Commission staff working document [CEC (2008b)], however, and given the track record of the two instruments so far, the preference of the Commission seems to have shifted away from establishing a uniform European TGC scheme in favor of creating an investor-friendly climate and optimizing existing national systems.[2]

Guaranteed FIT provide certainty about the achievable per-unit revenues from selling renewable electricity to the grid. While FIT have turned out to be very effective in countries such as Austria, Denmark, Germany and Spain, they cause market distortions to increase when electricity generation from renewables expands. In contrast, TGC are based on competitive market principles, typically featuring mandatory quota targets and certificate trading [e.g. Menanteau et al. (2003)]. Since TGC promise to enhance static and dynamic efficiency, they have attracted considerable attention. Over the years, they have been introduced in a number of countries with liberalized electricity markets [e.g. Berry and Jaccard (2001); Dinica and Arentsen (2003); Langniss and Wiser (2003); Lorenzoni (2003); Nielsen and Jeppesen (2003); Verbruggen (2004); Fan et al. (2005); Nishio and Asano (2006); Sáenz de Miera et al. (2008)]. More recently, the debate has been revolving around the interplay between TGC markets and markets for tradable CO_2 permits [e.g. Morthorst (2001); Jensen and Skytte (2003); Söderholm (2008)], and between TGC markets and liberalized power markets [e.g. Amundsen and Mortensen (2001, 2002); Jensen and Skytte (2002); Morthorst (2003); Amundsen and Bergman (2004)], respectively. Another active strand of research concerns financial risk of investors [Lemming (2003); Dinica (2006)].

While FIT is similar to a subsidy for suppliers of renewables, TGC constitute an internalization mechanism in the Baumol-Oates standard-price tradition [Baumol and Oates (1988)]. In fact, comparisons between taxes or subsidies and quota-based certificate schemes have so far been undertaken mainly in environmental economics, and in particular with regard to emission control. Denicolò (1999), for example, analyzes the effects of effluent charges and pollution permits when innovation is expected. Building on seminal work by Weitzman (1974, 1978), Pizer (1999a,b) studies the difference between a tax and quota policy under

[2]This apparent shift has to be seen in light of the very ambitious and binding policy target of achieving a 20% share of renewables in energy consumption by 2020. FIT, due to their relative simplicity in design, seem to find higher political acceptance and have become widespread in Europe and elsewhere in recent years. In CEC (2008a) it is reported that by 2007, of all 27 EU member countries, 18 had a FIT (or premium/bonus) system in place, seven a quota-based TGC system, and only two a tender system (Denmark for offshore wind, France for large projects).

uncertainty, finding that uncertainty causes the optimal amount of emission reduction to increase, which justifies a preference for taxation over quantity control. In the context of renewable energy, Madlener and Neustadt (2010) [see Chapter 7 of this dissertation] assess the impact of pre-commitment by government with respect to policy targets in the presence of cost-reducing innovation. In an empirical study, Palmer and Burtraw (2005) analyze the cost-effectiveness of two different renewable electricity policies (TGC vs. tax credits for renewable power production) in the U.S., and their impact on greenhouse gas emissions.

This paper is devoted to the issue of whether the diffusion of renewable power generating technologies can be better promoted by means of FIT or TGC, and in particular whether one of the schemes dominates the other in terms of cost-effectiveness and social welfare. We find that, given imperfectly competitive electricity markets, social welfare achieved under the optimal FIT policy is at least as high and likely to be strictly greater than social welfare under the optimal TGC policy, the latter importantly depending on the outcome of a strategic game in the market for tradable certificates. Our paper is organized as follows. Section 6.2 introduces the basic models used for contrasting effects of TGC and FIT in perfectly and imperfectly competitive markets for power. Under perfect competition, the equivalence of TGC and FIT is shown. This equivalence does not hold in a duopoly with quasi-symmetric costs, as demonstrated in section 6.3. Section 6.4 contains an evaluation of the two policies in terms of social welfare. Section 6.5 discusses policy implications, and section 6.6 concludes.

6.2 Promoting renewable electricity in a competitive market

We start our analysis with the simplest case, assuming that in a perfectly competitive electricity market there are N firms with equal electricity generation costs. Let there be only two options to produce electricity, either from fossil/nuclear or renewable resources (solar, wind, hydro, biomass etc.), with the second referred to as 'green electricity'. We assume that generation costs of fossil/nuclear power are generally lower than those of green electricity. However, green electricity cannot only help to avoid negative externalities from fossil/nuclear power generation, but also yield positive externalities in the form of different kinds of socio-economic benefits (e.g. creation of new employment, local value-added and infrastructure, spillovers from R&D in innovative energy technologies and systems).[3] The fact that these externalities are not sufficiently taken into account in decisions regarding the type and level of electricity production and consumption may motivate policy interventions such as the introduction of FIT and TGC.

6.2.1 FIT as a subsidy policy

The term 'subsidy' here refers to a transfer paid by the government or electricity consumers to the suppliers of green electricity. Thus, producers receive a surcharge s per unit of green electricity.[4] Given a competitive market, a representative generator of power faces the following optimization problem,

$$\max_{x_{\mathrm{b}},x_{\mathrm{g}}} \left[px_{\mathrm{b}} + (p+s)x_{\mathrm{g}} - C_{\mathrm{b}}(x_{\mathrm{b}}) - C_{\mathrm{g}}(x_{\mathrm{g}})\right], \tag{6.1}$$

where x_{b} and x_{g} denote the amounts of electricity produced from fossil/nuclear ('brown') fuels and renewable ('green') energy sources, respectively, $C_{\mathrm{b}}(x_{\mathrm{b}})$ is the cost function for electricity produced from fossil/nuclear fuel, $C_{\mathrm{g}}(x_{\mathrm{g}})$ is the cost function for green electricity, and p denotes the average market price for electricity. For an interior solution, the f.o.c. are

$$p \quad -C'_{\mathrm{b}}[x^*_{\mathrm{b}}] = 0 \tag{6.2}$$

$$p \quad +s - C'_{\mathrm{g}}[x^*_{\mathrm{g}}] = 0. \tag{6.3}$$

Inserting (6.2) into (6.3), we find that in an optimum with $x_{\mathrm{b}} > 0$ and $x_{\mathrm{g}} > 0$, the government subsidy s (or negative tax) has to be equal to the (absolute) difference between the marginal costs of green and conventional electricity evaluated at the optimum, $C'_{\mathrm{g}}[x^*_{\mathrm{g}}]$ and $C'_{\mathrm{b}}[x^*_{\mathrm{b}}]$, with $C'_{\mathrm{g}}[x^*_{\mathrm{g}}] > C'_{\mathrm{b}}[x^*_{\mathrm{b}}]$. The economic intuition behind this result is that if $s > C'_{\mathrm{g}}[x^*_{\mathrm{g}}] - C'_{\mathrm{b}}[x^*_{\mathrm{b}}]$, all generators will supply green electricity only; in contrast, if $s < C'_{\mathrm{g}}[x^*_{\mathrm{g}}] - C'_{\mathrm{b}}[x^*_{\mathrm{b}}]$, then no green electricity at all will be provided.

[3] Note that the use of green electricity may also lead to non-negligible negative externalities [e.g. Abbasi and Abbasi (2000); Tsoutsos et al. (2005)], but we assume here that these are generally smaller than the positive ones.

[4] In reality it is usually the power fed into the grid that counts, which due to on-site electricity consumption and transmission losses may be considerably less than gross production. This difference is neglected here for simplicity.

6.2.2 TGC as a quota-based policy

Rather than subsidizing green electricity, the government can also impose a green power production quota on each generator.[5] If a generator falls short of the quota, it faces a fine that increases with the shortfall. For each unit of green electricity produced, the generator obtains a certificate, providing proof of partial satisfaction of the norm.

Initially, assume that certificates are non-tradable. This assumption is natural given the assumption of identical costs across generators (no opportunity for trading). In section 6.3 below, the non-tradability assumption will be relaxed and a market for certificates introduced. For the situation of non-tradable certificates, the objective function that applies to a generator can be stated as:

$$\max_{x_{\mathrm{b}},x_{\mathrm{g}}} \left[p \cdot (x_{\mathrm{b}} + x_{\mathrm{g}}) - f \cdot (\bar{x}_{\mathrm{g}} - x_{\mathrm{g}}) - C_{\mathrm{b}}(x_{\mathrm{b}}) - C_{\mathrm{g}}(x_{\mathrm{g}}) \right], \tag{6.4}$$

where $\bar{x}_{\mathrm{g}}$ denotes the green electricity quota of the firm, f is the fine per unit of shortfall from the norm, and p, x_{b}, x_{g}, $C_{\mathrm{b}}(x_{\mathrm{b}})$, $C_{\mathrm{g}}(x_{\mathrm{g}})$ are the same as before. The f.o.c. with respect to x_{g} read

$$p \quad - C'_{\mathrm{b}}[x^*_{\mathrm{b}}] = 0 \tag{6.5}$$

$$p \quad + f - C'_{\mathrm{g}}[x^*_{\mathrm{g}}] = 0. \tag{6.6}$$

Note the similarity of (6.6) and (6.3). In fact the fine f plays the same role as the subsidy s, which therefore represents the shadow price of the quota. In an optimum, the unit price of the certificate should be equal to (slightly lower than) the value of the fine per unit.

[5] Note that in practice it is often the wholesalers or retailers, and sometimes even the final consumers of electricity, that are obliged to fulfil the quota.

6.2.3 Equivalence of FIT and TGC given identical costs

To show the equivalence of FIT and TGC, i.e. subsidy and quota-based policies, in terms of social welfare, we state the problem of a social planner as follows:[6]

$$W(Q, x_{\mathrm{g}}) = \max_{Q, x_{\mathrm{g}}} \int_0^Q p(\nu)d\nu - N \cdot C_{\mathrm{b}}\left(\frac{Q}{N} - x_{\mathrm{g}}\right) - N \cdot C_{\mathrm{g}}(x_{\mathrm{g}}) + E(Nx_{\mathrm{g}}), \tag{6.7}$$

where $Q = N(x_{\mathrm{b}} + x_{\mathrm{g}})$ stands for total electricity output produced by N firms, $p(\nu)$ for the inverse demand function, and $E(Nx_{\mathrm{g}})$ for the monetary value of the avoided negative and achieved positive externalities associated with green electricity production. As f.o.c. one obtains

$$p[Q^*] - C'_{\mathrm{b}}[x^*_{\mathrm{b}}] = 0 \tag{6.8}$$

$$C'_{\mathrm{b}}[x^*_{\mathrm{b}}] = C'_{\mathrm{g}}[x^*_{\mathrm{g}}] - E'[Nx^*_{\mathrm{g}}], \tag{6.9}$$

which determine the social optimum values of Q^*, x^*_{b} and x^*_{g}. Eq. (6.9) simply says that optimal aggregate output of green electricity must be such that the difference between the marginal cost and the marginal external benefit of green electricity is equal to the marginal cost of conventional power. If these quantities are known, the quota can be set as $\bar{x}_{\mathrm{g}} = x^*_{\mathrm{g}}$. The optimal subsidy level is given by $s^* = C'_{\mathrm{g}}[x^*_{\mathrm{g}}] - C'_{\mathrm{b}}[x^*_{\mathrm{b}}]$ from (6.3), and the optimal fine by $f^* = C'_{\mathrm{g}}[x^*_{\mathrm{g}}] - p$ from (6.6).

Obviously, subsidy and quota levels that are set according to the optimal values determined by maximizing social welfare will lead to the same level of green electricity production, yielding the same welfare. In this sense, and given our assumptions, the subsidy system and quota system are equivalent. Figure 6.1 illustrates the basic intuition behind these results. Let S^* denote the supply schedule reflecting that green power creates an external benefit to society. Therefore, it should be used at a rate $x^1_{\mathrm{g}} > x^0_{\mathrm{g}}$, with x^0_{g} being the outcome of supply S_0 based on private (marginal) cost and demand D. Clearly, the efficient quantity of green power can be attained by paying the optimal subsidy s^*, or imposing the optimal quota $\bar{x}^*_{\mathrm{g}}$.

[6] Seminal work on the equivalence of price and quantity control was provided by Bhagwati (1969) in the context of foreign trade (tariffs vs. quotas) and by Weitzman (1974) in the context of pollutant emission control (taxes vs. quotas), respectively.

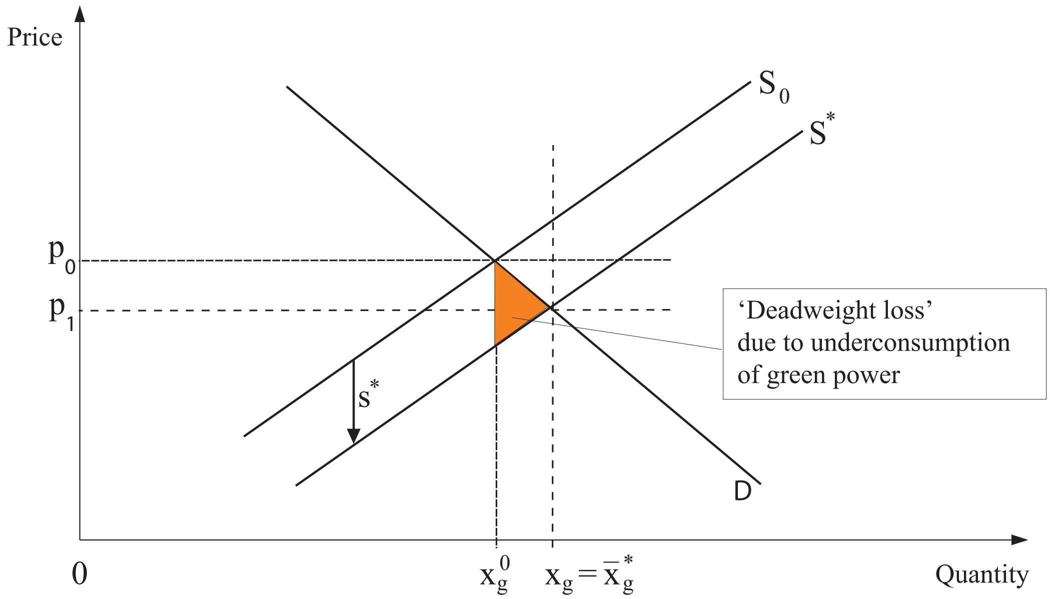


Figure 6.1: Equivalence of subsidy and quota-based policy given equal costs.

6.3 Duopoly market and quasi-symmetric costs

Studying the case of imperfectly competitive power markets as a duopoly game under quasi-symmetric costs can be justified on the following grounds. First, power markets are dominated by a few major players. For example, EdF still has a monopoly in France, PowerGen (now E.ON) has a market share of about 22 percent in the UK, and the four biggest suppliers in Germany, RWE, E.ON Energie, Vattenfall Europe and EnBW, together control more than two thirds of the market [cf. Bower et al. (2001); Matthes et al. (2005)]. Second, assuming the production costs of green power to be the same for all producers is not compatible with certificate trading. Therefore, we extend the basic model to the case of heterogeneous production costs in order to derive the potential for trade of green certificates.

Assume there are two generators in the market, firm 1 and 2, that have identical technology and hence cost functions in using fossil/nuclear fuel but different costs of generating renewable electricity. In this sense, the firms are 'quasi-symmetric'. The main reason to expect heterogeneous cost structures for green power is that it does not constitute yet a mature technology like that based on fossil or nuclear fuels, where competition presumably has forced operators to adopt the least-cost alternative. Therefore, producers of green power are assumed to employ different technologies, have more or less favorable siting of plants, use energy resources of different qualities, and employ different vintage mixes of a given technology.

To keep our model simple and to avoid multiple equilibria, we assume the cost function in using fossil/nuclear fuel $C_{\mathrm{b}}(x_{\mathrm{b}}) = c_{\mathrm{b}} x_{\mathrm{b}}$ to be linear and the difference of marginal cost between firm 1 and firm 2, $C'_{1\mathrm{g}}(y) - C'_{2\mathrm{g}}(y)$, to be a positive constant. Without loss of generality, we assume $C_{1\mathrm{g}}(y) > C_{2\mathrm{g}}(y)$ for any $y > 0$. With some loss of generality, but considerable gain in simplicity, let the demand function take on the following form,

$$p(x_{1\mathrm{b}}, x_{1\mathrm{g}}, x_{2\mathrm{b}}, x_{2\mathrm{g}}) = a - x_{1\mathrm{b}} - x_{1\mathrm{g}} - x_{2\mathrm{b}} - x_{2\mathrm{g}}, \quad a > 0, \tag{6.10}$$

which implies that consumers' willingness to pay is the same for fossil/nuclear and green power.

We start with the subsidy policy, focussing on the Cournot solution because power markets have been characterized by an absence of the fierce price competition one would expect in a Bertrand world. Limited price competition may be the result of collusion [Newbery (2002)], a variant of which is to stick to Cournot strategies. Moreover, under certain circumstances (e.g., capacity constraints), Cournot strategies continue to be pursued even under Bertrand-type competition [Kreps and Scheinkman (1983)]. In such a market set-up, firm 1 (the leader) believes that firm 2 (the follower) will react to firm 1's choice of green power produced. Thus in equilibrium firm 1 will have chosen a higher production level than in the case of a Cournot equilibrium and, consequently, firm 2 a lower level.[7]

6.3.1 Effect of subsidy on equilibrium

In this section, we assume that the subsidy is uniform, failing to take the difference in cost into account; the case of a non-uniform subsidy is discussed in Madlener and Neustadt (2010) [see Chapter 7 of this dissertation]. Here, the two firms face the following decision problem,

$$\max_{x_{i\mathrm{b}}, x_{i\mathrm{g}}} (a - x_{i\mathrm{b}} - x_{i\mathrm{g}} - x_{j\mathrm{b}} - x_{j\mathrm{g}})(x_{i\mathrm{b}} + x_{i\mathrm{g}}) + s x_{i\mathrm{g}} - c_{\mathrm{b}} x_{i\mathrm{b}} - C_{i\mathrm{g}}(x_{i\mathrm{g}}), \tag{6.11}$$

[7] A Stackelberg variant of this model would be an interesting extension, which is beyond the scope of this paper, however.

where $i, j = 1, 2$, and $i \neq j$.

We assume that the subsidy s is exogenous to and equal across firms. Generalizing condition (6.3), one can distinguish three different cases for the subsidy level (denoted S1–S3).

Case S1: $s \leq C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b} < C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_\mathrm{b}$

If $s < C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b}$, it is obvious that no green electricity will be produced because the subsidy does not make up for the efficient producer's cost disadvantage. Accordingly, the standard Cournot solution to the game is [cf. Kreps (1990), p. 326],

$$x^*_{1\mathrm{b}} = x^*_{2\mathrm{b}} = \frac{a - c_\mathrm{b}}{3}; \qquad x^*_{1\mathrm{g}} = x^*_{2\mathrm{g}} = 0. \tag{6.12}$$

If $s = C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b}$, then generator 2 is indifferent between producing green electricity and fossil/nuclear electricity.

Case S2: $C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b} \leq s < C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_\mathrm{b}$

In this case, the subsidy makes up for the cost disadvantage of green power for generator 2, but fails to do so for the less efficient generator 1, who therefore refrains from producing green electricity. The Cournot solution remains the same (in the sense that total electricity output of each firm remains unchanged), as compared to the case of a uniform quota.

So if $s > C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b}$, then generator 2 switches to green electricity, i.e.,

$$x^*_{1\mathrm{b}} = \frac{a - 2c_\mathrm{b} + C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - s}{3} \tag{6.13}$$

$$x^*_{2\mathrm{g}} = \frac{a - 2C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] + c_\mathrm{b} + 2s}{3} \tag{6.14}$$

$$x^*_{1\mathrm{g}} = x^*_{2\mathrm{b}} = 0. \tag{6.15}$$

Note that $x^*_{1\mathrm{b}}$ in (6.13) and $x^*_{2\mathrm{g}}$ in (6.14) are larger than in (6.12).

Case S3: $s \geq C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_{\mathrm{b}}$

If $s > C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_{\mathrm{b}}$, then the subsidy overcompensates the cost disadvantage of green power even for the less efficient generator 1. Therefore, both firms produce green electricity only. Accordingly, the optimal solutions are now

$$x^*_{1\mathrm{g}} = \frac{a + C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - 2C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] + s}{3} \tag{6.16}$$

$$x^*_{2\mathrm{g}} = \frac{a + C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - 2C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] + s}{3} \tag{6.17}$$

$$x^*_{1\mathrm{b}} = x^*_{2\mathrm{b}} = 0. \tag{6.18}$$

In the limiting case where $s = C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_{\mathrm{b}}$, generator 1 is indifferent between producing green and fossil/nuclear power, while generator 2, being efficient in the production of green power, supplies green electricity only.

Optimal subsidy level

The results derived in the previous subsection show that the equilibrium solutions to the Cournot game strongly depend upon the level of the subsidy. This raises the issue of determining the optimal subsidy level. In analogy to (6.7), let social welfare be given by

$$\begin{aligned} W^j(Q, x_{1\mathrm{g}}, x_{2\mathrm{g}}; s) \;=\; & \int_0^Q p(\nu)d\nu - c_{\mathrm{b}}(Q - x_{1\mathrm{g}} - x_{1\mathrm{g}}) \\ & - C_{1\mathrm{g}}(x_{1\mathrm{g}}) - C_{2\mathrm{g}}(x_{2\mathrm{g}}) + E(x_{\mathrm{g}}), \end{aligned} \tag{6.19}$$

with W^j denoting the social welfare gains associated with case j ($j = 1$, 2, and 3) of Section 6.3.1. We assume that in case 2, s is slightly greater than $C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_{\mathrm{b}}$, and in case 3 slightly greater than $C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_{\mathrm{b}}$, in order to avoid ambiguity.

To facilitate comparison between the cases, the externality function associated with green electricity takes the form $E(x_{\mathrm{g}}) = \beta x_{\mathrm{g}}$, $\beta > 0$. While it would certainly be interesting to elaborate on possible alternative functional forms of $E(x_{\mathrm{g}})$, and consequences for the outcome, such an analysis is beyond the scope of this paper and saved for future research.

The parameter β (called 'welfare parameter' henceforth) implies a constant marginal social benefit from producing green electricity. Using the equilibrium values given in (6.12) to (6.18), the welfare associated with the three cases can be written as follows:

$$W^1 = \left(a - \frac{Q}{2}\right) Q - c_\mathrm{b} Q \tag{6.20}$$

$$W^2 = \left(a - \frac{Q}{2}\right) Q + \beta x_{2\mathrm{g}} - c_\mathrm{b}(Q - x_{2\mathrm{g}}) - C_{2\mathrm{g}}(x_{2\mathrm{g}}) \tag{6.21}$$

$$W^3 = \left(a - \frac{Q}{2}\right) Q + \beta Q - C_{1\mathrm{g}}(x_{1\mathrm{g}}) - C_{2\mathrm{g}}(x_{2\mathrm{g}}). \tag{6.22}$$

As is to be expected, whether or not the welfare parameter β exceeds the marginal cost parameters is of crucial importance. For $\beta > C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_\mathrm{b}$, the welfare parameter is larger than the additional costs incurred by firm 1, so that it is optimal if both firms produce green electricity. Conversely, if the positive externality $\beta x_{2\mathrm{b}}$ exceeds the extra costs of producing green electricity for firm 2, it is optimal if firm 2 produces green instead of brown electricity.

More specifically, we can distinguish the following situations:

(A) if $\beta > C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_\mathrm{b}$, then $W^3 > W^2 > W^1$. Hence the optimal subsidy is the lower bound of the subsidy interval in case 3, i.e., $s^*_\mathrm{A} = C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_\mathrm{b}$.

(B) if $\beta = C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_\mathrm{b}$, then $W^3 = W^2 > W^1$. The welfare gains remain the same for $s^*_\mathrm{B} = C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_\mathrm{b}$ and $s^{**}_\mathrm{B} = C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b}$, though the amounts of green electricity produced are different.

(C) if $C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b} \leq \beta < C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_\mathrm{b}$, then $W^2 > W^3$ and $W^2 \geq W^1$. The optimal subsidy is thus equal to the lower bound of the subsidy interval in case 2, i.e., $s^*_\mathrm{C} = C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b}$.

(D) if $\beta < C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b}$, then $W^1 > W^2 > W^3$. Therefore, the optimal subsidy is zero, because none of the rates are effective in promoting green power.

Figure 6.2 summarizes the optimal subsidy schedule for different values of the welfare parameter β.

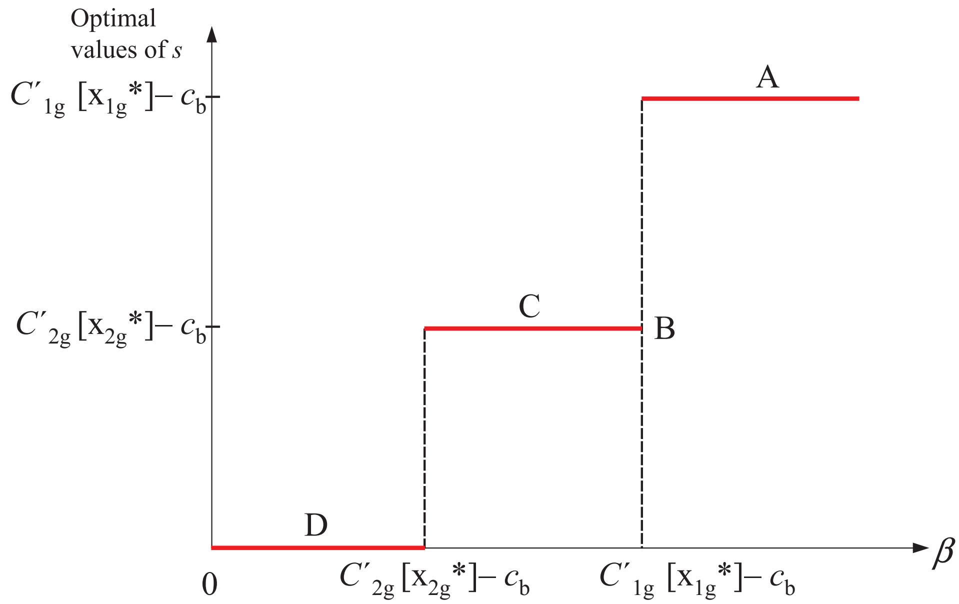


Figure 6.2: Optimal subsidy levels vs welfare parameter β of green electricity, cases A through D.

6.3.2 Quota-based policy

Building on (6.4) of section 6.2.2, the decision problem faced by the two firms in a duopoly market can be written as

$$\begin{aligned} \max_{x_{i\mathrm{b}}, x_{i\mathrm{g}}} & [(a - x_{i\mathrm{b}} - x_{i\mathrm{g}} - x_{j\mathrm{b}} - x_{j\mathrm{g}})(x_{i\mathrm{b}} + x_{i\mathrm{g}}) + z(\tilde{x}_{i\mathrm{g}} - \tilde{x}_{j\mathrm{g}}) \\ & - f(\bar{x}_{\mathrm{g}} - x_{i\mathrm{g}} - \tilde{x}_{i\mathrm{g}}) - c_{\mathrm{b}} x_{i\mathrm{b}} - C_{i\mathrm{g}}(x_{i\mathrm{g}})], \\ \text{s.t.} \quad & x_{i\mathrm{g}} + x_{j\mathrm{g}} = 2\bar{x}, \end{aligned} \tag{6.23}$$

with $\tilde{x}_{i\mathrm{g}}$ ($\tilde{x}_{j\mathrm{g}}$) denoting the amount of certificates sold (purchased), respectively, $i, j = 1$ or 2, and $i \neq j$, f denoting the fine per unit as in (6.4), and z denoting the certificate price. Note that the constraint implies that the amount of green certificates produced by the two firms must not exceed the industry quota – i.e. we assume that once the quota is satisfied the certificate price drops to zero. Thus, there is no incentive to produce more green electricity than is required by the quota target.

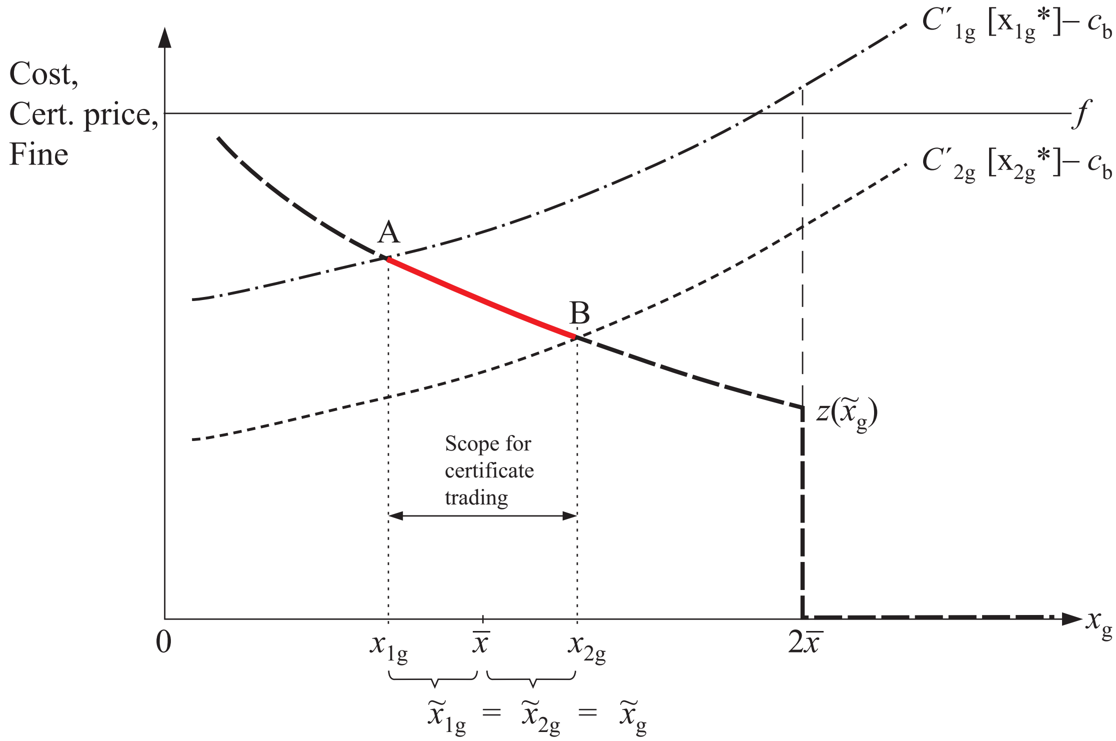


Figure 6.3: Effects of the TGC policy in a duopoly game

In the above model, each firm has two choice variables. However, given the assumption that the difference $C'_{1g}(y) - C'_{2g}(y)$ is a positive constant for any $y > 0$, generator 1's choice of x_{1g} boils down to a choice between 0 and $\bar{x}_g$, depending on the ordering of f, z, and the level of the difference in marginal costs of producing green and brown electricity (cf. figure 6.3). If it is in the interest of generator 1 to purchase green electricity certificates at all, it must also be (at least weakly) in its interest to go all the way. Therefore, we can find a Nash equilibrium by comparing the firms' payoffs for $x^*_{1g} \in \{0, \bar{x}_g\}$, which will be shown in the following section. Before turning to the Nash equilibrium, however, we briefly discuss the different possible cases, of which we will examine the two that are desirable from a social welfare point of view.

First, the fine could fall short of the difference in marginal costs for firms 1 and 2, in which both firms prefer to pay the fine and produce conventional electricity only.

Second, if the fine is larger than the cost difference for firm 2 but lower than the marginal cost difference for firm 1 of producing green instead of brown electricity, then we can distinguish two sub-cases: either the certificate price is larger or smaller than the fine. In the former case, firm 1 prefers to pay the fine, rather than buying certificates from firm 2, while in the latter case it would buy certificates from firm 2 up to its quota (provided, of course, that trading is possible).

Third, the fine is set at a higher level than the difference in the marginal costs of producing green instead of conventional electricity for both generators. In this case we can again distinguish two sub-cases, one where z is lower than the cost difference for firm 1 (so that it has an incentive to buy certificates up to $\bar{x}$ from generator 2) and a situation where it is higher (in which case generator 1 would self-generate green electricity up to its quota[8]).

In the following, we consider the two cases where certificate trading occurs (case I) and the case where firm 1 self-produces green certificates (case II).

6.3.3 Nash equilibrium under the quota-based policy

We now elaborate the Nash equilibrium for the quota-based policy, by comparing the firms' payoffs under the two socially desirable strategies stated at the end of the previous section 6.3.2. Hence, in the discussion that follows, we distinguish the two cases I and II.

Case I: $x^*_{1\mathrm{g}} = 0$

Case I refers to a situation where the cost difference for firm 1 of producing green instead of brown electricity is lower than the fine f but higher than the certificate price z. Therefore, it is cheaper for firm 1 to buy certificates from firm 2. Given that $x^*_{1\mathrm{g}} = 0$, generator 2 is required to produce at least $2\bar{x}_\mathrm{g}$ units of green electricity in order to satisfy the industry quota. This can be summarized as follows,

$$x^*_{1\mathrm{g}} = 0 \Leftrightarrow f > C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_\mathrm{b} > z > C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b} \tag{6.24}$$

or as

$$C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_\mathrm{b} > f > z > C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b}.$$

[8] Note that in this situation generator 1 would possibly be forced at some point to leave the market, as its costs of producing green electricity are too high.

Firm 1 solves the problem

$$\max_{x_{1b}, x_{1g}} \Pi_1 = (a - x_{1b} - x_{2b} - x_{2g})x_{1b} - c_b x_{1b} - z\tilde{x}_g - \max\{0, f(\bar{x}_g - \tilde{x}_{1g})\}, \tag{6.25}$$

with f.o.c.:

$$\frac{\partial \Pi_1}{\partial x_{1b}} = a - 2x^*_{1b} - x^*_{2b} - x^*_{2g} - c_b = 0. \tag{6.26}$$

Observing the constraint $x_{2g} \geq 2\bar{x}_g$, firm 2 solves

$$\begin{aligned} \max_{x_{2b}, x_{2g}} L(x_{2b}, x_{2g}, \lambda) \quad &= \quad (a - x_{1b} - x_{2b} - x_{2g})(x_{2b} + x_{2g}) - c_b x_{2b} + z\tilde{x}_g \\ &\quad - \max\{0, f(\bar{x}_g - x_{2g} + \tilde{x}_g)\} - C_{2g}(x_{2g}) + \lambda(2\bar{x}_g - x_{2g}), \end{aligned}$$

with $\lambda \geq 0$ denoting the Lagrange multiplier. The f.o.c. read,

$$\frac{\partial L}{\partial x_{2b}} \quad = \quad a - x^*_{1b} - 2x^*_{2b} - 2x^*_{2g} - c_b = 0 \tag{6.27}$$

$$\frac{\partial L}{\partial x_{2g}} \quad = \quad a - x^*_{1b} - 2x^*_{2b} - 2x^*_{2g} - c_b - C'_{2g}(x^*_{2g}) + f + \lambda = 0. \tag{6.28}$$

From (6.26) and (6.27) we get $x^*_{1b} = (a - c_b)/3$, and from (6.27) and (6.28) we obtain $x^*_{2b} + x^*_{2g} = (a - c_b)/3$.

Given that firm 1 does not produce any green electricity, we need to distinguish two subcases. In the first subcase (Ia, Appendix 6.7), firm 2 self-generates its own quota, while in the second subcase (Ib, Appendix 6.7) it produces twice the individual firm's quota (and hence is able to sell the excess certificates to firm 1, which faces higher production costs).

Eqs. (6.26) and (6.27) say that the two firms will produce the same total quantity of electricity, determined by the maximum possible market demand and the cost of producing electricity from fossil/nuclear fuel. From (6.27) and (6.28) we obtain

$$\lambda = C'_{2g}[x^*_{2g}] - c_b - f. \tag{6.29}$$

Eq. (6.29) indicates that if $C'_{2g}[x^*_{2g}] - c_b > f$, such that $\lambda > 0$, generator 2 will only produce green electricity up to the industry quota as required by our assumptions, due to the Kuhn–Tucker condition. Note that trading of certificates is also possible as long as $C'_{1g}[x^*_{1g}] - c_b \geq f$.

However, if $C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b} = f$ (and hence $\lambda = 0$), generator 2 has an incentive to produce at least the quota required from the industry. Also note that different values of $f \in [C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b}, C'_{1\mathrm{g}}[x^*_{1\mathrm{g}}] - c_\mathrm{b}]$ only affect the distribution of profits between the two firms, with no impact on the amount of certificate trading and social welfare. Therefore, we first focus on the case $C'_{2\mathrm{g}}[x^*_{2\mathrm{g}}] - c_\mathrm{b} = f$ as a benchmark.

The optimal quota continues to be determined as in eqs. (6.7)–(6.9), except that $C'_\mathrm{b}[x^*_\mathrm{b}] = c_\mathrm{b}$. Hence $\bar{x}_\mathrm{g} = x^*_\mathrm{g}/2$ still holds. As long as $2\bar{x}_\mathrm{g} \leq (a - c_\mathrm{b})/3$ [see eq. (6.40)], generator 2 produces $(a - c_\mathrm{b})/3 - 2\bar{x}_\mathrm{g}$ units of electricity using fossil/nuclear fuel and $2\bar{x}_\mathrm{g}$ units of green electricity. As to $2\bar{x}_\mathrm{g} > (a - c_\mathrm{b})/3$, recall that a denotes the willingness to pay for the first kWh of electricity, while c_b symbolizes the (constant) marginal cost of fossil/nuclear power, which makes $a - c_\mathrm{b}$ a very large number. It is unlikely for $\bar{x}_\mathrm{g}$ to exceed one sixth of that number, justifying that this case is neglected.

So far, we have assumed that generator 1 is the only buyer of generator 2's extra certificates. However, there may be another agent willing to purchase the certificates at the market price, for example, an environmental protection agency or a foundation promoting renewable energy. Since the equilibrium price of certificates is determined in such a manner that generator 2 is indifferent between producing green or fossil/nuclear fuel electricity, the presence of an additional bidder might cause generator 2 to produce green electricity in excess of the quota. However, this would make the system a combination of quantity and price policies. The reason is that these extra purchases, resulting in an increase of the value of the certificates, can be viewed as a subsidy. It is possible that such a policy mix is more effective in promoting green power than either one of the two policy instruments individually. However, a detailed analysis of such a mixed policy is beyond the scope of this paper.

Case II: $x^*_{1\mathrm{g}} = \bar{x}_\mathrm{g}$

We now turn to the case of generator 1 producing green electricity to satisfy the quota. This is possible if the difference in the marginal cost of producing green and brown electricity is strictly lower than the certificate price (and the fine), or formally,

$$x^*_{1\mathrm{g}} = \bar{x}_\mathrm{g} \Leftrightarrow f > z > C'_{1\mathrm{g}}[x_{1\mathrm{g}}] - c_\mathrm{b}. \tag{6.30}$$

With no external agent purchasing, the condition $x_{1\mathrm{g}} = \bar{x}_{\mathrm{g}}$ or $x_{2\mathrm{g}} = \bar{x}_{\mathrm{g}}$ continues to hold. Firm 1's optimization problem now reads,

$$\max_{x_{1\mathrm{b}},x_{1\mathrm{g}}} \Pi_1 = (a - x_{1\mathrm{b}} - x_{2\mathrm{b}} - \bar{x}_{\mathrm{g}} - x_{2\mathrm{g}})(x_{1\mathrm{b}} + \bar{x}_{\mathrm{g}}) - c_{\mathrm{b}}x_{1\mathrm{b}} - C_{1\mathrm{g}}(\bar{x}_{\mathrm{g}}),$$

with f.o.c.,

$$\frac{\partial \Pi_1}{\partial x_{1\mathrm{b}}} = a - 2x^*_{1\mathrm{b}} - 2\bar{x}_{\mathrm{g}} - x^*_{2\mathrm{b}} - x_{2\mathrm{g}} - c_{\mathrm{b}} = 0; \tag{6.31}$$

while firm 2 solves the problem

$$\begin{aligned} \max_{x_{2\mathrm{b}},x_{2\mathrm{g}}} \Pi_2 &= (a - x_{1\mathrm{b}} - x_{2\mathrm{b}} - \bar{x}_{\mathrm{g}} - x_{2\mathrm{g}})(x_{2\mathrm{b}} + x_{2\mathrm{g}}) - c_{\mathrm{b}}x_{2\mathrm{b}} - C_{2\mathrm{g}}(x_{2\mathrm{g}}) \\ &\quad - \max\{0, f(\bar{x}_{\mathrm{g}} - x_{2\mathrm{g}})\}. \end{aligned} \tag{6.32}$$

Since $x^*_{2\mathrm{g}}$ can only take on $\bar{x}_{\mathrm{g}}$ or $2\bar{x}_{\mathrm{g}}$ as optimal values, we concentrate on $x^*_{2\mathrm{g}} \geq \bar{x}_{\mathrm{g}}$ and only consider the first-order condition concerning variable $x_{2\mathrm{b}}$, which reads:

$$\frac{\partial \Pi_2}{\partial x_{2\mathrm{b}}} = a - 2x^*_{2\mathrm{b}} - 2x_{2\mathrm{g}} - x^*_{1\mathrm{b}} - \bar{x}_{\mathrm{g}} - c_{\mathrm{b}} = 0. \tag{6.33}$$

From (6.31) and (6.33) we get $x^*_{1\mathrm{b}} = (a - c)/3 - \bar{x}_{\mathrm{g}}$ and $x^*_{2\mathrm{b}} = (a - c)/3 - x_{2\mathrm{g}}$.

As before, we have to distinguish two subcases (IIa and IIb in Appendix 6.7), in both of which firm 1 self-generates green electricity up to its individual quota, while firm 2 either produces $\bar{x}_{\mathrm{g}}$ or twice $\bar{x}_{\mathrm{g}}$.

		Firm 2: $\bar{x}_{\mathrm{g}}$	Firm 2: $2\bar{x}_{\mathrm{g}}$
Firm 1	0	$-f\bar{x}_{\mathrm{g}},\ -c_{\mathrm{b}}\bar{x}_{\mathrm{g}} - C_{2\mathrm{g}}(\bar{x}_{\mathrm{g}})$	$-z\bar{x}_{\mathrm{g}},\ (z - 2c_{\mathrm{b}})\bar{x}_{\mathrm{g}} - C_{2\mathrm{g}}(2\bar{x}_{\mathrm{g}})$
Firm 1	$\bar{x}_{\mathrm{g}}$	$c_{\mathrm{b}}\bar{x}_{\mathrm{g}} - C_{1\mathrm{g}}(\bar{x}_{\mathrm{g}}),\ c_{\mathrm{b}}\bar{x}_g - C_{2\mathrm{g}}(\bar{x}_g)$	$c_{\mathrm{b}}\bar{x}_{\mathrm{g}} - C_{1\mathrm{g}}(\bar{x}_{\mathrm{g}}),\ 2c_{\mathrm{b}}\bar{x}_{\mathrm{g}} - C_{2\mathrm{g}}(2\bar{x}_{\mathrm{g}})$

Figure 6.4: Payoffs to producers of green electricity under the TGC policy

If firm 2's strategy is to produce the minimal quota, then firm 1's best response is $x_{1\mathrm{g}} = \bar{x}_{\mathrm{g}}$ if $C_{1\mathrm{g}}(\bar{x}_{\mathrm{g}}) - c_{\mathrm{b}}\bar{x}_{\mathrm{g}} \leq f\bar{x}_{\mathrm{g}}$. However, if firm 2's strategy is to produce twice the quota, then firm 1's best response is $x_{1\mathrm{g}} = 0$ if $C_{1\mathrm{g}}(\bar{x}_{\mathrm{g}}) - c_{\mathrm{b}}\bar{x}_{\mathrm{g}} \geq z\bar{x}_{\mathrm{g}}$. If firm 1's strategy is to produce 0, then firm 2's best response is $2\bar{x}_{\rm g}$, provided the following condition holds: $C_{2\rm g}(2\bar{x}_{\rm g}) - C_{2\rm g}(\bar{x}_{\rm g}) \leq (z - c_{\rm b})\bar{x}_{\rm g}$. This condition holds due to the assumptions made in this paper. If firm 1's strategy is to produce the minimal quota, then firm 2's best response is $\bar{x}_{\rm g}$, provided that the following condition is satisfied: $C_{2\rm g}(2\bar{x}_{\rm g}) - C_{2\rm g}(\bar{x}_{\rm g}) \geq c_{\rm b}\bar{x}_{\rm g}$. This condition again holds due to the assumptions made. Therefore, if the condition $z\bar{x}_{\rm g} \leq C_{1\rm g}(\bar{x}_{\rm g}) - c_{\rm b}\bar{x}_{\rm g} \leq f\bar{x}_{\rm g}$ is satisfied, this game has two Nash equilibria in pure strategies: $(0, 2\bar{x}_{\rm g})$ and $(\bar{x}_{\rm g}, \bar{x}_{\rm g})$.

6.4 Welfare comparison between subsidy and quota-based policies

In spite of the simplifying assumptions made, a welfare comparison between a price-subsidy and a quota policy may be worthwhile because it promises to provide some guidance to policy-makers regarding the choice of instruments for promoting renewable energy use.

6.4.1 Welfare gains under the subsidy policy

Since our main interest is to discuss how to efficiently promote green power, case S1 (section 6.3.1) can be disregarded since it is fossil/nuclear only. In addition, case S3 (section 6.3.1) is not realistic because it predicts that all firms exclusively produce green power, which would presuppose extremely high green electricity quota. Therefore, we only examine the case associated with condition $C'_{2\rm g}[x^*_{2\rm g}] - c_{\rm b} \leq \beta < C'_{1\rm g}[x^*_{1\rm g}] - c_{\rm b}$, i.e. case S2 of section 6.3.1. The pertinent welfare function is repeated from (6.21) for convenience,

$$W^s = \left(a - \frac{Q}{2}\right) Q - c_{\rm b}(Q - x_{2\rm g}) - C_{2\rm g}(x_{2\rm g}) + \beta x_{2\rm g}, \tag{6.34}$$

where Q continues to be total production of both types of electricity. Remember that in case S2 we have $x^*_{1\rm g} = x^*_{2\rm b} = 0$ and the optimal subsidy is given by $s^*_C = C'_{2\rm g}[x^*_{2\rm g}] - c_{\rm b}$, thus the total production given the optimal subsidy scheme can be determined as $Q^s = x^s_{1\rm b} + x^s_{2\rm g} = 2(a - c_{\rm b})/3$, with $x^s_{2\rm g} = (a - c_{\rm b})/3$ denoting the amount of green energy produced by firm 2. Therefore, social welfare achieved by the optimal subsidy scheme is

$$W^s = (Q^s)^2 - C_{2\rm g}\left(\frac{Q^s}{2}\right) + \, (c_{\rm b} + \beta)\,\frac{Q^s}{2}. \tag{6.35}$$

6.4.2 Welfare gains under the quota-based policy

If marginal costs of green power are increasing, the optimal quota cannot be determined directly. To match the production of green electricity in the subsidy case, we simply assume that $\bar{x}_\mathrm{g}$ is equal to $(a - c_\mathrm{b})/6$, which may constitute a rather frequent solution [see the discussion after eq. (6.29) in section 6.3.3]. The welfare function for the quota-based certificate system can then be specified as

$$W^q = \left(a - \frac{Q}{2}\right) Q - c_\mathrm{b}(Q - x_{2\mathrm{g}}) - C_{2\mathrm{g}}(x_{2\mathrm{g}}) + \beta x_{2\mathrm{g}}. \tag{6.36}$$

The total amount of energy $Q^q = 2(a - c_\mathrm{b})/3$ produced given the quota-based policy is identical with Q^s. However, w.r.t. to $x_{2\mathrm{g}}$, we have to distinguish the following two possible pure Nash equilibrium outcomes of the game described in section 6.3.3:

(i) Welfare achieved in the Nash equilibrium $(x^{q1}_{1\mathrm{g}}, x^{q1}_{2\mathrm{g}}) = (0, 2\bar{x}_\mathrm{g})$,

$$W^{q1} = (Q^q)^2 - C_{2\mathrm{g}}\left(\frac{Q^q}{2}\right) + (c_\mathrm{b} + \beta)\,\frac{Q^q}{2}. \tag{6.37}$$

Since $Q^q = Q^s$, the welfare under the quota-based policy realized in this Nash equilibrium is equal to the welfare under the optimal subsidy policy.

(ii) Welfare achieved in the Nash equilibrium $(x^{q2}_{1\mathrm{g}}, x^{q2}_{2\mathrm{g}}) = (\bar{x}_\mathrm{g}, \bar{x}_\mathrm{g})$,

$$W^{q2} = (Q^q)^2 - C_{2\mathrm{g}}\left(\frac{Q^q}{4}\right) + (c_\mathrm{b} + \beta)\,\frac{Q^q}{4}. \tag{6.38}$$

Note that the welfare level W^{q2} is lower than $W^{q1} = W^s$ if

$$C_{2\mathrm{g}}(2y) - C_{2\mathrm{g}}(y) < (c_b + \beta)y \quad \text{for } y > 0.$$

This condition is satisfied if the level of marginal social benefit of green electricity β is sufficiently high. Hence, the subsidy policy guarantees a welfare level which might not be achieved with the quota policy if the Nash equilibrium $(\bar{x}_\mathrm{g}, \bar{x}_\mathrm{g})$ is played. A comparison of both firms' profits as well as of their sum shows that only firm 1 is better off in the socially efficient equilibrium $(0, 2\bar{x}_\mathrm{g})$. Under our assumptions, both firm 2's profit and total producer surplus are likely to be higher in the socially less desirable equilibrium $(\bar{x}_\mathrm{g}, \bar{x}_\mathrm{g})$. Therefore, even if the Cournot game were repeated an infinite number of times, no cooperative equilibrium would occur with both firms choosing the socially desirable strategies.

6.4.3 Welfare of subsidy and quota-based policies in a quasi-symmetric duopoly

Comparing welfare levels given in (6.35), (6.37), and (6.38), one sees that, for sufficiently high marginal social benefits of green energy, $W^s = W^{q1} > W^{q2}$. This result implies that even with imperfect competition and quasi-heterogeneous costs, subsidies should be preferred to tradable certificates. This is intuitive, since outcomes in a duopoly crucially depend on whether firms pursue price- or quantity-oriented strategies and FIT could be said to be price-oriented whereas TGC is quantity-oriented. However, the results established above given imperfectly competitive markets seem to hinge on two crucial assumptions. The first is that TGC are tradable. This means that price is the signal to competitors, precisely as the subsidy under FIT. And since the quota and the subsidy are set as to optimally internalize the externalities present, the information content of price is the same under both regimes. Second, competitors pursue optimal duopoly strategies regardless of the choice of internalization policy adopted by the government.

In addition to these two basic premises, there are simplifying assumptions that should be kept in mind. Specifically, the cost of administering subsidies and/or quota are neglected and therefore assumed equal. However, when it comes to start-up costs, a certificate system may require more resources than a subsidy system, especially for establishing appropriate regulation and regulatory control. Also, information regarding cost and marginal revenues on the part of competitors as well as marginal positive externalities of green power on the part of government was assumed perfect. Yet due to cost heterogeneity, the amount of information required for calculating the optimal subsidy typically increases with a growing number of firms. Although the setting of the optimal quota requires similar information, the heterogeneity of firms does not enter their determination, causing it to be relatively straightforward and hence probably less costly than a subsidy system. These considerations also suggest that a generalization from duopoly to oligopoly would be straightforward.

Further, we found that subsidies provide more incentives for green power precisely when its marginal social benefits are high (as in case S3 of section 6.3.1). A pure quota-based TGC system lacks this feature.

6.5 Policy implications

Based on several models incorporating imperfect competitiveness of markets for power for added realism, we find that the subsidy (FIT) approach, when implemented at its socially optimal value, leads to a welfare gain, which is not necessarily achieved with the quota-based (TGC) policy. Furthermore, the subsidy policy is generally preferred by utilities, likely because it does not call into question their right to cause a certain amount of pollution when using fossil/nuclear fuel input. At the same time, subsidies do provide stronger incentives for pollution-abating innovation than quotas by directly favoring production of green electricity. Since the future of green electricity importantly depends on future technological progress for lowering its cost of production, subsidies are also more efficient dynamically.

On the other hand, the financing of subsidies requires tax revenue. When the (economic or political) cost of additional taxation is high, like in the United States (but also in Scandinavian countries e.g.), the quota-based approach may provide a viable alternative. As found in the present analysis, tradable green certificates are more efficient than non-tradable ones regardless of market structure. Trade in certificates is likely to develop because green power does not yet rely on a mature cost-minimizing technology, contrary to fossil/nuclear generation. Moreover, since the cost of running a market for certificates is lower once the market is established, the disadvantage of the quota-based policy will gradually wane, without however reaching the dynamic efficiency of the subsidy approach.

6.6 Conclusions

This paper starts from the suspicion that the conventional wisdom, claiming a tax/subsidy (FIT) and a quota/certificate (TGC) policy scheme to be equivalent in terms of static efficiency, might not hold if markets for power are imperfectly competitive. Based on a duopoly model in which the two competitors differ in terms of their marginal cost of producing 'green' power, we show that, if both schemes are implemented at their respective socially optimal values, the subsidy policy is at least equivalent, but can be superior to the quota policy depending on the outcome of the game in the market for green certificates. The subsidy and the tradable certificate contain the same (correct) price information entering competitors' strategy choices, which are of the Cournot type regardless of the scheme considered. Interestingly, however, only one of two pure-strategy Nash equilibria under the quota-based policy corresponds to the unique equilibrium outcome under the subsidy policy whereas the other Nash equilibrium (in which both firms produce green energy) leads to a lower welfare level. In view of the technological heterogeneity of green power generation, it is important that certificates are tradable. Also, the possible equivalence breaks down as soon as incentives for pollution-abating innovation are considered as well. Thus, the subsidy is the preferable approach; on the other hand, its financing may meet with a high marginal cost of taxation.

Acknowledgements

The authors acknowledge comments received from Eberhard Feess, Kaushik Deb, Michael Kuenzle and Julia Meyer, as well as participants in the 2006 Annual Meeting of the Swiss Society of Economics and Statistics (9-10 March 2006, Lugano, Switzerland) on an earlier version of the manuscript.

6.7 Appendix

Subcase Ia: $x^*_{1\mathrm{b}} = \dfrac{a - c_\mathrm{b}}{3}$; $x^*_{1\mathrm{g}} = 0$; $x^*_{2\mathrm{b}} = \dfrac{a - c_\mathrm{b}}{3} - \bar{x}_\mathrm{g}$; $x^*_{2\mathrm{g}} = \bar{x}_\mathrm{g}$

Market demand:

$$a - x_{1\mathrm{b}} - x_{2\mathrm{b}} - x_{2\mathrm{g}} = \frac{a + 2c_\mathrm{b}}{3} \tag{6.39}$$

Firms' profits:

$$\begin{aligned}
\Pi_1 &= \frac{a + 2c_\mathrm{b}}{3}\,\frac{a - c_\mathrm{b}}{3} - c_\mathrm{b} \cdot \frac{a - c_\mathrm{b}}{3} - f\bar{x}_\mathrm{g} = \frac{(a - c_\mathrm{b})^2}{9} - f\bar{x}_\mathrm{g}; \\
\Pi_2 &= \frac{a + 2c_\mathrm{b}}{3}\,\frac{a - c_\mathrm{b}}{3} - c_\mathrm{b} \cdot \left(\frac{a - c_\mathrm{b}}{3} - \bar{x}_\mathrm{g}\right) + z\bar{x}_\mathrm{g} - C_{2\mathrm{g}}(\bar{x}_\mathrm{g}) - f\bar{x}_\mathrm{g} \\
&= \frac{(a - c_\mathrm{b})^2}{9} - c_\mathrm{b}\bar{x}_\mathrm{g} - C_{2\mathrm{g}}(\bar{x}_\mathrm{g})
\end{aligned}$$

Subcase Ib: $x_{1\mathrm{b}}^* = \dfrac{a-c_\mathrm{b}}{3}$; $x_{1\mathrm{g}}^* = 0$; $x_{2\mathrm{b}}^* = \dfrac{a-c_\mathrm{b}}{3} - 2\bar{x}_\mathrm{g}$; $x_{2\mathrm{g}}^* = 2\bar{x}_\mathrm{g}$

Market demand:

$$a - x_{1\mathrm{b}} - x_{2\mathrm{b}} - x_{2\mathrm{g}} = \frac{a+2c_\mathrm{b}}{3} \tag{6.40}$$

Firms' profits:

$$\begin{aligned}
\Pi_1 &= \frac{a+2c_\mathrm{b}}{3}\,\frac{a-c_\mathrm{b}}{3} - c_\mathrm{b}\frac{a-c_\mathrm{b}}{3} - z\bar{x}_\mathrm{g} \;=\; \frac{(a-c_\mathrm{b})^2}{9} - z\bar{x}_\mathrm{g};\\
\Pi_2 &= \frac{a+2c_\mathrm{b}}{3}\,\frac{a-c_\mathrm{b}}{3} - c_\mathrm{b}\left(\frac{a-c_\mathrm{b}}{3} - 2\bar{x}_\mathrm{g}\right) + z\bar{x}_\mathrm{g} - C_{2\mathrm{g}}(2\bar{x}_\mathrm{g}) - f\bar{x}_\mathrm{g}\\
&= \frac{(a-c_\mathrm{b})^2}{9} + (z-2c_\mathrm{b})\bar{x}_\mathrm{g} - C_{2\mathrm{g}}(2\bar{x}_\mathrm{g})
\end{aligned}$$

Subcase IIa (symmetry): $x_{1\mathrm{b}}^* = x_{2\mathrm{b}}^* = \dfrac{a-c_\mathrm{b}}{3} - \bar{x}_\mathrm{g}$; $x_{1\mathrm{g}}^* = x_{2\mathrm{g}}^* = \bar{x}_\mathrm{g}$

Market demand:

$$a - x_{1\mathrm{b}} - x_{1\mathrm{g}} - x_{2\mathrm{b}} - x_{2\mathrm{g}} = \frac{a+2c_\mathrm{b}}{3} \tag{6.41}$$

Firms' profits:

$$\begin{aligned}
\Pi_1 &= \frac{a+2c_\mathrm{b}}{3}\,\frac{a-c_\mathrm{b}}{3} - c_\mathrm{b}\left(\frac{a-c_\mathrm{b}}{3} - \bar{x}_g\right) - C_{1\mathrm{g}}(\bar{x}_\mathrm{g}) \;=\; \frac{(a-c_\mathrm{b})^2}{9} + c_\mathrm{b}\bar{x}_\mathrm{g} - C_{1\mathrm{g}}(\bar{x}_\mathrm{g});\\
\Pi_2 &= \frac{a+2c_\mathrm{b}}{3}\,\frac{a-c_\mathrm{b}}{3} - c_\mathrm{b}\left(\frac{a-c_\mathrm{b}}{3} - \bar{x}_\mathrm{g}\right) - C_{2\mathrm{g}}(\bar{x}_\mathrm{g}) \;=\; \frac{(a-c_\mathrm{b})^2}{9} + c_\mathrm{b}\bar{x}_\mathrm{g} - C_{2\mathrm{g}}(\bar{x}_\mathrm{g})
\end{aligned}$$

Subcase IIb: $x_{1\mathrm{b}}^* = \dfrac{a-c_\mathrm{b}}{3} - \bar{x}_\mathrm{g}$; $x_{1\mathrm{g}}^* = \bar{x}_\mathrm{g}$; $x_{2\mathrm{b}}^* = \dfrac{a-c_\mathrm{b}}{3} - 2\bar{x}_\mathrm{g}$; $x_{2\mathrm{g}}^* = 2\bar{x}_\mathrm{g}$

Market demand:

$$a - x_{1\mathrm{b}} - x_{1\mathrm{g}} - x_{2\mathrm{b}} - x_{2\mathrm{g}} = \frac{a+2c_\mathrm{b}}{3} \tag{6.42}$$

Firms' profits:

$$\begin{aligned}
\Pi_1 &= \frac{a+2c_\mathrm{b}}{3}\,\frac{a-c_\mathrm{b}}{3} - c_\mathrm{b}\left(\frac{a-c_\mathrm{b}}{3} - \bar{x}_\mathrm{g}\right) - C_{1\mathrm{g}}(\bar{x}_\mathrm{g}) \;=\; \frac{(a-c_\mathrm{b})^2}{9} + c_\mathrm{b}\bar{x}_\mathrm{g} - C_{1\mathrm{g}}(\bar{x}_\mathrm{g});\\
\Pi_2 &= \frac{a+2c_\mathrm{b}}{3}\,\frac{a-c_\mathrm{b}}{3} - c_\mathrm{b}\left(\frac{a-c_\mathrm{b}}{3} - 2\bar{x}_\mathrm{g}\right) - C_{2\mathrm{g}}(2\bar{x}_\mathrm{g}) \;=\; \frac{(a-c_\mathrm{b})^2}{9} + 2c_\mathrm{b}\bar{x}_\mathrm{g} - C_{2\mathrm{g}}(2\bar{x}_\mathrm{g})
\end{aligned}$$

Renewable energy policy in the presence of innovation: does government pre-commitment matter?

Reinhard Madlener and Ilja Neustadt

Abstract: In a perfectly competitive market with a possibility of technological innovations we contrast guaranteed feed-in tariffs for electricity from renewables and tradable green certificates from a dynamic efficiency and social welfare point of view. Specifically, we model decisions about the technological innovation within the framework of a game-theoretic model and discuss implications for optimal policy design under different assumptions regarding regulatory pre-commitment. We find that for the case of technological innovation subsidy policies are preferable. Further, in terms of dynamic efficiency, no pre-commitment policies are shown to be at least as good as the pre-commitment ones.

Keywords: Renewable electricity, feed-in tariffs, regulatory pre-commitment, tradable green certificates, quota target, innovation, energy policy
JEL classification: Q42, Q48

Chapter 7

Renewable energy policy in the presence of innovation: does government pre-commitment matter?

7.1 Introduction

Renewable energy is considered an important element in a sustainable energy development. In many countries renewable energy promotion policies have been put into place. As far as electricity generation from renewables is concerned, there has been much debate in recent years about the relative merits of guaranteed feed-in tariffs (FIT) and tradable green certificates (TGC), mainly in the form of qualitative discussion (e.g. Menanteau et al., 2003; Nielsen and Jeppesen, 2003; Berry, 2002), and much less so in the form of more rigorous formal analysis (e.g. Amundsen and Mortensen, 2001).

Building on seminal work by Weitzman (1974, 1978), Pizer (1999a,b) studies the non-equivalence of tax and quota policies given uncertainty and shows that uncertainty causes the optimal amount of emission abatement to increase, which justifies a preference for price over quantity control. Madlener et al. (2009) show that in terms of static efficiency a price (subsidy) policy to promote renewable energy is equivalent to a quantity (quota) policy for a competitive but not generally a duopoly market for power when competitors have different production costs for renewable (but not conventional) energy. In this paper, we extend the static analysis to incorporate technological innovation that lowers the (increasing) marginal cost of production of electricity from renewable sources.

From environmental economics it is known that the dynamic efficiency of a policy depends on whether or not the government pre-commits to a certain policy target (e.g. Denicolò, 1999). In our analysis we want to find out which of the two policy instruments provides a stronger incentive for innovation favoring renewable or "green" electricity in two cases, (1) when the government adjusts its policy in response to innovation (no pre-commitment), and (2) when it cannot react immediately to innovation (pre-commitment). In contrast to Denicolò (1999), we find that the relative merits of the subsidy and quota policies are the same in the two scenarios from the point of view of social welfare maximization. However, in terms of dynamic efficiency, this equivalence does not necessarily hold. Rather, the no pre-commitment policy is shown to support equilibrium outcomes with innovations that might not be attainable under pre-commitment.

The remainder of the paper is organized as follows. Section 2 derives optimal subsidy and quota policies for assuming no pre-commitment on the part of the government when innovation is present. Section 3 contains the analogous analysis for the pre-commitment case. Section 4 discusses the results obtained in Sections 2 and 3 and concludes.

7.2 Optimal policy in the presence of innovation: no pre-commitment case

In the no pre-commitment case, the government is assumed to have the information, ability and obligation to respond to technological innovation by adjusting its subsidy or quota policy, respectively. Let there be $N+1$ competitive electricity generators in the market, one of them being the potential innovator, assumed to possess the patent covering the rights for the new technology. Innovation reduces the marginal cost of green electricity, and the innovator can license the new technology to other producers in return of a royalty. Let us assume that prior to innovation all firms have an identical cost structure for producing green electricity of the simplistic form

$$C_{\mathrm{g}}(x_{\mathrm{g}}) = b_1 x_{\mathrm{g}} + b_2 {x_{\mathrm{g}}}^2, \tag{7.1}$$

with $b_1 > 0$, $b_2 > 0$, to reflect decreasing marginal returns (DMR) in the production of green electricity. DMR is a sensible assumption because the use of renewables (in particular solar and wind) involves technologies that have not yet reached maturity. Accordingly, there is scope for (exogenous) innovation, resulting in a new cost function of the form

$$C_{\mathrm{gn}}(x_{\mathrm{g}}) = b_{1\mathrm{n}} x_{\mathrm{g}} + b_2 {x_{\mathrm{g}}}^2, \tag{7.2}$$

where C_{gn} denotes the cost function after innovation and $b_{1\mathrm{n}} < b_1$ the reduced part of the marginal cost. Note that b_2 is unaffected by the innovation for simplicity ($b_{1\mathrm{n}} < b_1$ is sufficient to mitigate DMR). Thus, $(b_1 - b_{1n})$ reflects the importance of the innovation. The cost function for brown electricity (i.e. from conventional sources such as coal, nuclear etc.) is assumed to be linear, $C_{\mathrm{b}}(x_{\mathrm{b}}) = c_{\mathrm{b}} x_{\mathrm{b}}$.

The R&D investment required for the innovation is denoted by $R[C_{\mathrm{g}}(x_{\mathrm{g}}) - C_{\mathrm{gn}}(x_{\mathrm{g}})]$, with $R'(\cdot) > 0$ and $R''(\cdot) > 0$. This means that the R&D outlay increases progressively as a function of the size of the achievable cost reduction. Therefore, R&D does not display increasing marginal returns, reflecting the fact that no particular technology has dominated the market for renewable electricity to this day. Given the continuity assumptions made in (7.1) and (7.2), for any fixed value of x_{g}, $R[C_{\mathrm{g}}(x_{\mathrm{g}}) - C_{\mathrm{gn}}(x_{\mathrm{g}})]$ can be rewritten as $R(b_1 - b_{1\mathrm{n}})$. We consider a parametric version of function $R(\cdot)$ of the form $R(b_1 - b_{1\mathrm{n}}) = r(b_1 - b_{1\mathrm{n}})^2$, with parameter $r > 0$ reflecting the concavity of the function. In particular, the higher r, the higher the marginal cost of innovation.

On the demand side, we assume that brown and green electricity are perfect substitutes. Thus the demand function for electricity takes the following linear form:

$$p(Q) = a - Q = a - \sum_{i=1}^{N+1} \left(x_{\mathrm{ib}} + x_{\mathrm{ig}}\right),$$

where Q denotes the total quantity of electricity supplied in the market, x_{ib}, the quantity of conventional electricity produced by firm i, and x_{ig}, that of green electricity. Further, we assume that $b_1 < c_b$, i.e. marginal costs of green electricity are lower than those of brown electricity for small quantities, and $(c_b - b_{1n})$ is sufficiently smaller than $b_2(a - c_b)$, i.e. the average electricity price on the market, p, will always be given by the marginal cost of brown electricity c_b.

The government observes whether a firm operates with the old or the new technology[1] and is assumed to maximize social welfare. The externality function of green electricity[2] (including avoided social cost of producing brown electricity) is assumed to have a simple, linear-quadratic form:

$$D(x_{\mathrm{g}}) = d_1 x_{\mathrm{g}} - d_2 x_{\mathrm{g}}^2, \qquad d_1,\, d_2 > 0. \tag{7.3}$$

The quadratic term reflects the fact that marginal avoided social cost of brown electricity decreases with higher quantities of green electricity produced and might attain negative values if large quantities of green electricity are produced.[3] In order to exclude the possibility of extremely high social cost of additional production of green power, we additionally assume that parameter d_2 is sufficiently small such that $d_2(N+1)(c_{\mathrm{b}} - b_{1\mathrm{n}}) < b_2 d_1$.

7.2.1 Subsidy policy

Subsidy (or negative tax) here refers to a transfer paid by the government or electricity consumers to the suppliers of green electricity. Thus, producers receive a surcharge s per unit of green electricity.[4] The decisions of the agents can be represented by a game with the following players: firms 1, 2, ..., $N+1$, and government G. Without loss of generality, let us assume that firm no. 1 is the potential innovator.

[1] This is a plausible assumption since, in reality, the electricity producers are required to file the technical description of their power generating technology to the regulator.

[2] Note that, in the real world, the quantification of the (positive and negative) externalities associated with power generation from renewables is subject to several complications (e.g. Söderholm and Sundqvist, 2003). The value of the external benefits (including avoided environmental damages and learning-by-doing effects) is likely to depend on the particular composition of the technology portfolio used to produce electricity, and thus also the amount of the brown electricity displaced and the (environmental) benefit incurred.

[3] This can be motivated by arguing that with more intensive utilization of renewables, environmentally and socially less benign projects are also being realized.

[4] In reality it is usually the power fed into the grid that counts, which due to on-site electricity consumption and transmission losses may be considerably less than gross production. This difference is neglected here for simplicity.

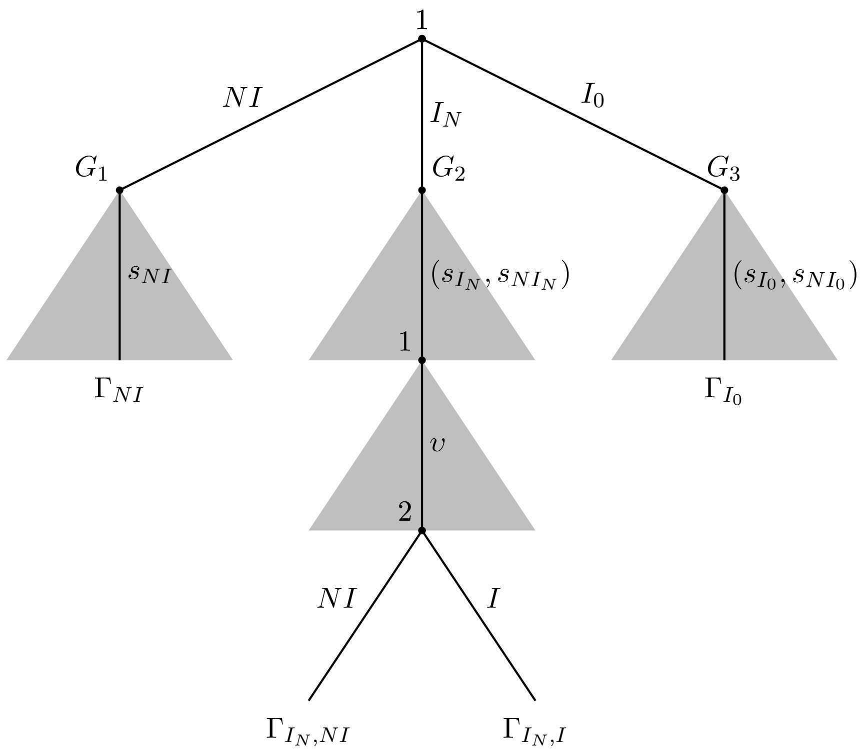


Figure 7.1: Extensive-form game representation, no pre-commitment case, subsidy policy

Now we analyze the decision sequencing under subsidy control with no pre-commitment. There are three decision stages, described in the following and summarized in Figure 7.1.

Stage I. Firm 1 decides either not to innovate (NI), to innovate and offer N licenses in the competition stage III (I_N), or to innovate and offer no licenses in stage III (I_0).

Stage II. Given the decision of firm 1 in stage I, the government determines the subsidy levels for non-innovating and innovating firms in order to maximize social welfare.

(IIa) If firm 1 did not innovate, the government introduces a subsidy s_{NI} per unit of output for all firms (decision node G_1).

(IIb) If firm 1 did innovate and announced to offer N licenses in stage III, the government introduces two levels of subsidy: s_{I_N} for the innovator and the firms that adopted the new technology and s_{NI_N} for the firms that did not adopt the new technology (decision node G_2).

(IIc) Finally, if firm 1 did innovate but announced that it will offer no licenses in stage III, the subsidies are s_{I_0} for the innovator and s_{NI_0} for the competitors (decision node G_3).

Stage III. Given the innovation decision of firm 1 and the decision of the government about the subsidy level, firms 1, 2, …, $N+1$ compete in quantities.

(IIIa) If firm 1 did not innovate, all firms have identical cost functions $C_{\mathrm{g}}(\cdot)$ and compete in quantities given subsidy level s_{NI} per unit of green electricity (subgame Γ_{NI}).

(IIIb) If firm 1 did innovate and committed to offer N licenses in stage III, then it first offers licenses to N competitors in return of a royalty v given subsidy levels s_{I_N} and s_{NI_N}. Firms $2, \ldots, N+1$ can either accept (I) or reject (NI) this offer. Since firms $2, \ldots, N+1$ are identical, we assume that either all of them will reject the offer and operate with cost function $C_{\mathrm{g}}(\cdot)$ (competition in quantities will take place in subgame $\Gamma_{I_N,NI}$) or all of them will accept it and operate with cost function $C_{\mathrm{gn}}(\cdot)$ (competition in subgame $\Gamma_{I_N,I}$).

(IIIc) If firm 1 did innovate but announced that it will offer no licenses in stage III, then firm 1, operating with cost function $C_{\mathrm{gn}}(\cdot)$, and firms $2, \ldots, N+1$, operating with cost function $C_{\mathrm{g}}(\cdot)$, respectively, compete in quantities given their subsidy levels s_{I_0} and s_{NI_0} (subgame Γ_{I_0}).

These three decision stages define an extensive-form game as shown in Figure 7.1. The information revealed in the earlier stages of this game is taken as given in the corresponding subsequent stages. Thus, in the earlier stages, rational players anticipate the equilibrium outcomes in every subsequent stage. Each game branch starting with an information set can thus be considered as a subgame, giving rise to the Subgame Perfect Equilibrium (SPE) as the solution concept to be applied. As usual, the SPE solution can be obtained by backward induction.

Lemma 7.2.1. *In subgame Γ_{NI} (stage IIIa), all firms' quantities of green electricity are given by*

$$x_{i\mathrm{g}}(NI, s_{NI}) \quad = \quad \frac{c_{\mathrm{b}} - b_1 + s_{NI}}{2b_2}. \tag{7.4}$$

Proof: see Appendix on p.180.

Lemma 7.2.2. *In stage IIIb, firm 1's equilibrium offer υ^* is given by*

$$\upsilon^* = \begin{cases} (b_1 - b_{1\mathrm{n}}) - (s_{NI_N} - s_{I_N}) & \textit{if } (b_1 - b_{1\mathrm{n}}) - (s_{NI_N} - s_{I_N}) < \dfrac{c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_N}}{2}; \\ \dfrac{c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_N}}{2} & \textit{otherwise.} \end{cases}$$

This offer is always accepted by a firm of type 2 in equilibrium[5]*. Quantities of green electricity produced by firm 1 and firms of type 2 are*

$$\begin{aligned} x_{1\mathrm{g}}(I_N, (s_{NI_N}, s_{I_N})) &= \frac{c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_N}}{2b_2}; \\ x_{2\mathrm{g}}(I_N, (s_{NI_N}, s_{I_N})) &= \begin{cases} \dfrac{(c_b - b_1 + s_{NI_N})}{2b_2} & \textit{if } (b_1 - b_{1\mathrm{n}}) - (s_{NI_N} - s_{I_N}) < \dfrac{c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_N}}{2}; \\ \dfrac{c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_N}}{4b_2} & \textit{otherwise.} \end{cases} \end{aligned}$$

Proof: see Appendix on p.180.

Lemma 7.2.3. *In subgame Γ_{I_0} (stage IIIc), quantities of green electricity produced by firm 1 and firms of type 2 are given by*

$$\begin{aligned} x_{1\mathrm{g}}(I_0, (s_{NI_0}, s_{I_0})) &= \frac{c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_0}}{2b_2}; \\ x_{2\mathrm{g}}(I_0, (s_{NI_0}, s_{I_0})) &= \frac{c_\mathrm{b} - b_1 + s_{NI_0}}{2b_2}. \end{aligned}$$

Proof: see Appendix on p.182.

Lemma 7.2.4. *In stage IIa (subgame starting at node G_1), the government chooses subsidy level*

$$s^*_{NI} = \frac{b_2 d_1 - d_2(N+1)(c_\mathrm{b} - b_1)}{b_2 + d_2(N+1)}.$$

Proof: see Appendix on p.183.

Lemma 7.2.5. *In stage IIb (subgame starting at node G_2), the government chooses any combination of subsidy levels*

$$(s^*_{NI_N}, s^*_{I_N}) = \left(s^*_{NI_N}, \frac{[2b_2 N - d_2(N+2)^2](c_\mathrm{b} - b_{1\mathrm{n}}) + 2(N+2)b_2 d_1}{d_2(N+2)^2 + 4b_2} \right),$$

[5] As usual, we assume that in the case of indifference firms of type 2 decide in favor of the adoption of the new technology.

where

$$s^*_{NI_N} \geq (b_1 - b_{1\mathrm{n}}) + \frac{[b_2(N-2) - d_2(N+2)^2](c_{\mathrm{b}} - b_{1\mathrm{n}}) + (N+2)b_2 d_1}{d_2(N+2)^2 + 4b_2}.$$

Proof: see Appendix on p.183.

Lemma 7.2.6. *In stage IIc (subgame starting at node G_3), the government chooses subsidy levels*

$$s^*_{NI_0} = s^*_{I_0} = \frac{b_2 d_1 - d_2(N+1)(c_{\mathrm{b}} - b_1) - d_2(b_1 - b_{1\mathrm{n}})}{b_2 + d_2(N+1)}.$$

Proof: see Appendix on p.185.

Proposition 7.2.7. *The subgame-perfect equilibrium strategies on the equilibrium path of the innovation game with subsidy control and no pre-commitment policy are given as follows. Firm 1 does not innovate (NI) if $(r-1)(\Delta b_1)^2 + 2\beta\Delta b_1 - \alpha\beta^2 \leq 0$ where*

$$\begin{aligned} \alpha &= \frac{b_2}{4}\left(\frac{2(N+2)^3}{[4b_2 + d_2(N+2)^2]^2} - \frac{1}{[b_2 + d_2(N+1)]^2}\right) > 0; \\ \beta &= c_{\mathrm{b}} - b_{1\mathrm{n}} + d_1 > 0; \\ \Delta b_1 &= b_1 - b_{1\mathrm{n}} \end{aligned}$$

and innovates and offers N licenses (I_N) otherwise. The royalty and quantities in equilibrium are given by

$$\upsilon^* = \frac{b_2(N+2)(c_{\mathrm{b}} - b_{1\mathrm{n}} + d_1)}{d_2(N+2)^2 + 4b_2};$$

$$x^*_{1\mathrm{g}}(NI) = \frac{c_{\mathrm{b}} - b_1 + d_1}{2[b_2 + d_2(N+1)]};\ x^*_{1\mathrm{g}}(I_N) = \frac{(N+2)(c_{\mathrm{b}} - b_{1\mathrm{n}} + d_1)}{d_2(N+2)^2 + 4b_2}.$$

Government sets subsidy levels

$$\begin{aligned} s^*_{NI} &= \frac{b_2 d_1 - d_2(N+1)(c_{\mathrm{b}} - b_1)}{b_2 + d_2(N+1)}, \\ s^*_{NI_N} &\in \left\{ s :\ s \leq \frac{[b_2(N-2) - d_2(N+2)^2](c_{\mathrm{b}} - b_1 + d_1) + b_2(N+2)(b_1 - b_{1\mathrm{n}})}{4b_2 + d_2(N+2)^2} + d_1 \right\}, \\ s^*_{I_N} &= \frac{[2b_2 N - d_2(N+2)^2](c_{\mathrm{b}} - b_{1\mathrm{n}}) + 2(N+2)b_2 d_1}{d_2(N+2)^2 + 4b_2}. \end{aligned}$$

Firms of type 2 innovate (I) if firm 1 chooses I_N and produce quantities

$$\begin{aligned} x^*_{2\mathrm{g}}(NI) &= \frac{c_\mathrm{b} - b_1 + d_1}{2[b_2 + d_2(N+1)]}, \\ x^*_{2\mathrm{g}}(I_N) &= \frac{(N+2)(c_\mathrm{b} - b_{1\mathrm{n}} + d_1)}{2[d_2(N+2)^2 + 4b_2]}. \end{aligned}$$

Proof: see Appendix on p.185.

7.2.2 Quota-based policy

Instead of subsidizing green electricity, the government can also impose a quota target for green power on each generator.[6] For each unit of green electricity produced, the firm receives a certificate providing evidence of partial satisfaction of the target imposed[7]. If a firm falls short of achieving the quota target, it faces a fine f that increases with the shortfall (cf. Madlener et al., 2009).

As with the subsidy-based policy, we consider an extensive-form game with the following structure. There are three decision stages.

Stage I. Firm 1 decides either not to innovate (NI), to innovate and offer N licenses in the competition stage III (I_N), or to innovate and offer no royalties in stage III (I_0).

Stage II. Given the decision of firm 1 in stage I, the government determines the quotas to be satisfied and the fines for firms falling short of the quota for non-innovating and innovating firms, in order to maximize social welfare.

(IIa) If firm 1 did not innovate, the government introduces a quota $\bar{x}_{NI}$ and a fine f_{NI} per unit of output falling short of the quota for all firms (decision node G_1).

(IIb) If firm 1 did innovate and announced to offer N licenses in stage III, the government introduces two pairs of quotas and fines: $(\bar{x}_{I_N}, f_{I_N})$ for the innovator and those firms that adopted the new technology and $(\bar{x}_{NI_N}, f_{NI_N})$ for those firms that did not adopt the new technology (G_2).

[6] In practice it is often the wholesalers or retailers, and sometimes even the final consumers of electricity, that are obligated to fulfil the quota target.

[7] Admittedly, the assumption that the market for tradable certificates is perfectly competitive and efficient may, especially in poorly designed or managed schemes, be quite a strong one (e.g. Nilsson and Sundqvist, 2007; Söderholm, 2008).

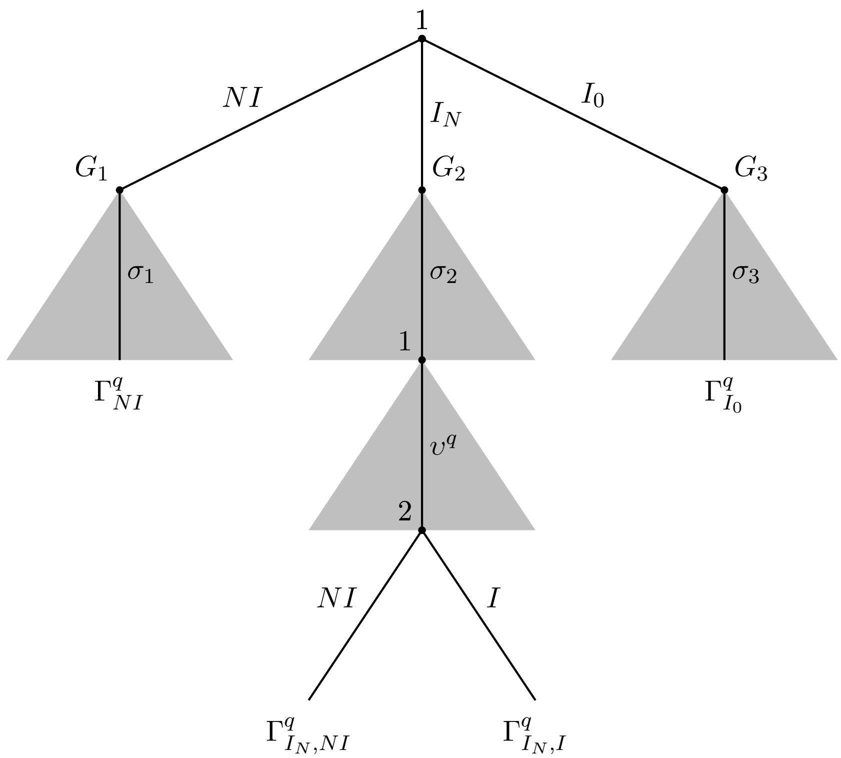


Figure 7.2: Extensive-form game representation, no pre-commitment case, quota policy. The actions of the government σ_i for each node G_i, $i = 1, 2, 3$, are defined as follows: $\sigma_1 = (\bar{x}_{NI}, f_{NI})$, $\sigma_2 = ((\bar{x}_{NI_N}, f_{NI_N}), (\bar{x}_{I_N}, f_{I_N}))$, $\sigma_3 = ((\bar{x}_{I_0}, f_{I_0}), (\bar{x}_{NI_0}, f_{NI_0}))$.

(IIc) Finally, if firm 1 did innovate but announced that it will offer no licenses in stage III, the quotas and fines set by the government are $(\bar{x}_{I_0}, f_{I_0})$ for the innovator and $(\bar{x}_{NI_0}, f_{NI_0})$ for the competitors (G_3).

Stage III. Given the innovation decision of firm 1 and the decision of the government about the quotas and fines, firms 1, 2, ..., $N+1$ compete in quantities.

(IIIa) If firm 1 did not innovate, all firms have an identical cost function $C_{\mathrm{g}}(\cdot)$ and compete in quantities given the quota and fine levels $(\bar{x}_{NI}, f_{NI})$ (subgame Γ^q_{NI}).

(IIIb) If firm 1 did innovate and announced to offer N licenses in stage III, then firm 1 first offers licenses to N competitors for a royalty υ^q, given the quota and the fine levels $(\bar{x}_{I_N}, f_{I_N})$, $(\bar{x}_{NI_N}, f_{NI_N})$. Firms 2, 3, ..., $N+1$ (firms of type 2) can accept or reject this offer. Since firms of type 2 are identical, we assume that either all of them will reject the offer and operate with the new cost function $C_{\mathrm{g}}(\cdot)$ (competition in quantities will take place in subgame $\Gamma^q_{I_N,NI}$) or all of them will accept it and operate with cost function $C_{\text{gn}}(\cdot)$ (competition in subgame $\Gamma^q_{I_N,I}$).

(IIIc) If firm 1 did innovate but announced that it will offer no licenses in stage III, then firm 1, operating with cost function $C_{\text{gn}}(\cdot)$, and firms of type 2, operating with cost function $C_{\text{g}}(\cdot)$, compete in quantities, given their quota and fine levels $(\bar{x}_{I_0}, f_{I_0})$, $(\bar{x}_{NI_0}, f_{NI_0})$ (subgame $\Gamma^q_{I_0}$).

These three decision stages define an extensive-form game as shown in Figure 7.2. Like in the subsidy case, we apply the solution concept of the Subgame Perfect Equilibrium (SPE).

Lemma 7.2.8. *In stage IIIa (subgame Γ^q_{NI}), all firms produce quantity*

$$x_{ig}(NI, f_{NI}) = \frac{c_{\text{b}} - b_1 + f_{NI}}{2b_2}.$$

Proof: see Appendix on p.186.

Lemma 7.2.9. *In stage IIIb, firm 1's equilibrium offer υ^{q*} is given by*

$$\upsilon^{*q} = \begin{cases} \upsilon^{q\max} & \text{if } \upsilon^{q\max} < \dfrac{c_{\text{b}} - b_{1\text{n}} + f_{I_N}}{2}; \\ \dfrac{c_{\text{b}} - b_{1\text{n}} + f_{I_N}}{2} & \text{otherwise.} \end{cases}$$

where

$$\upsilon^{q\max} = \sqrt{(c_{\text{b}} - b_1 + f_{I_N})^2 + 4b_2(f_{NI_N}\bar{x}_{NI_N} - f_{I_N}\bar{x}_{I_N})} - (c_{\text{b}} - b_{1\text{n}} + f_{I_N}).$$

This offer is always accepted by firms of type 2 in the equilibrium[8]. Firm 1 produces quantity

$$x_{1\text{g}}(I_N, (f_{NI_N}, f_{I_N}), \upsilon^q) = \frac{c_{\text{b}} - b_{1\text{n}} + f_{I_N}}{2b_2}.$$

[8] By assumption, firms of type 2 adopt the new technology if indifferent.

The quantity of green electricity produced by any firm of type 2, $x_{2\mathrm{g}}(I_N, (f_{NI_N}, f_{I_N}, \bar{x}_{I_N}, \bar{x}_{NI_N}))$, amounts to

$$\begin{cases} \dfrac{2(c_b - b_1 + f_{NI_N}) - \sqrt{(c_\mathrm{b} - b_1 + f_{I_N})^2 + 4b_2(f_{NI_N}\bar{x}_{NI_N} - f_{I_N}\bar{x}_{I_N})}}{2b_2} & \text{if } \upsilon^{\mathrm{qmax}} < \dfrac{c_\mathrm{b} - b_{1\mathrm{n}} + f_{I_N}}{2}; \\ \dfrac{c_\mathrm{b} - b_{1\mathrm{n}} + f_{I_N}}{4b_2} & \text{otherwise.} \end{cases}$$

Proof: see Appendix on p.186.

Lemma 7.2.10. *In stage IIIb (subgame $\Gamma^q_{I_0}$), quantities of green electricity produced by firm 1 and firms of type 2, respectively, are given by*

$$\begin{aligned} x_{1\mathrm{g}}(I_0, f_{NI_0}) &= \frac{c_\mathrm{b} - b_{1\mathrm{n}} + f_{I_0}}{2b_2}; \\ x_{2\mathrm{g}}(I_0, f_{NI_0}) &= \frac{c_\mathrm{b} - b_1 + f_{NI_0}}{2b_2}. \end{aligned}$$

Proof: see Appendix on p.188.

Lemma 7.2.11. *In stage IIa (subgame starting at node G_1), the government chooses fine level*

$$f^*_{NI} = \frac{b_2 d_1 - d_2(N+1)(c_\mathrm{b} - b_1)}{b_2 + d_2(N+1)},$$

while the quota level $\bar{x}_{NI}$ can be deliberately set by the government.

Proof: see Appendix on p.188.

Lemma 7.2.12. *In stage IIb (subgame starting at node G_2), the optimal decision of the government is given by any combination of fines*

$$(f^*_{NI_N}, f^*_{I_N}) = \left(f^*_{NI_N}, \frac{[2b_2 N - d_2(N+2)^2](c_\mathrm{b} - b_{1\mathrm{n}}) + 2(N+2)b_2 d_1}{d_2(N+2)^2 + 4b_2} \right)$$

where $f^*_{NI_N}$ satisfies inequality

$$\sqrt{(c_\mathrm{b} - b_1 + f^*_{I_N})^2 + 4b_2(f^*_{NI_N}\bar{x}_{NI_N} - f^*_{I_N}\bar{x}_{I_N})} \;\geq\; \frac{3}{2}(c_\mathrm{b} - b_{1\mathrm{n}} + f^*_{I_N}). \tag{7.5}$$

The government's choice of quotas $\bar{x}_{NI_N}, \bar{x}_{I_N}$ is constrained by inequality (7.5).

Proof: see Appendix on p.189.

Lemma 7.2.13. *In stage IIc (subgame starting at node G_3), government chooses fine levels*

$$f^*_{NI_0} = f^*_{I_0} = \frac{b_2 d_1 - d_2(N+1)(c_\mathrm{b} - b_1) - d_2(b_1 - b_{1\mathrm{n}})}{b_2 + d_2(N+1)}.$$

Proof: see Appendix on p.190.

Proposition 7.2.14. *The subgame-perfect equilibrium strategies on the equilibrium path of this game are given as follows. Firm 1 does not innovate (NI) if* $\Delta b_1 \in \left(0, \frac{\sqrt{B^2-4AC}-B}{2A}\right]$, *where*

$$\begin{aligned}
A &= \frac{b_2}{4[b_2 + d_2(N+1)]^2 + r} > 0;\\
B &= \frac{d_2(N+1)}{b_2 + d_2(N+1)}\bar{x}_{NI} - \frac{2b_2}{4[b_2 + d_2(N+1)]^2\beta};\\
C &= -\alpha\beta^2 - \left(\frac{d_2(N+1)}{b_2 + d_2(N+1)} + d_1\right)\bar{x}_{NI} - \left(\frac{d_2(N+2)^2 - 2b_2 N}{4b_2 + d_2(N+2)^2}\beta + d_1\right)\bar{x}_I < 0
\end{aligned}$$

with

$$\begin{aligned}
\alpha &= \frac{b_2}{4}\left(\frac{2(N+2)^3}{[4b_2 + d_2(N+2)^2]^2} - \frac{1}{[b_2 + d_2(N+1)]^2}\right) > 0;\\
\beta &= c_\mathrm{b} - b_{1\mathrm{n}} + d_1 > 0,
\end{aligned}$$

and innovates and offers N licenses (I_N) otherwise. It offers N licenses in return of a royalty

$$v^{*q} = \frac{b_2(N+2)(c_\mathrm{b} - b_{1\mathrm{n}} + d_1)}{d_2(N+2)^2 + 4b_2}$$

and produces quantities of green electricity

$$\begin{aligned}
x^*_{1\mathrm{g}}(NI) &= \frac{c_\mathrm{b} - b_1 + d_1}{2[b_2 + d_2(N+1)]},\\
x^*_{1\mathrm{g}}(I_N) &= \frac{(N+2)(c_\mathrm{b} - b_{1\mathrm{n}} + d_1)}{d_2(N+2)^2 + 4b_2}.
\end{aligned}$$

Government chooses fine levels

$$
\begin{aligned}
f^*_{NI} &= \frac{b_2 d_1 - d_2(N+1)(c_\mathrm{b} - b_1)}{b_2 + d_2(N+1)}, \\
f^*_{NI_N} &\geq \frac{\left\{[2b_2 N - (N+2)^2 d_2](c_\mathrm{b} - b_{1\mathrm{n}}) + 2(N+2)b_2 d_1\right\} \bar{x}_I}{\bar{x}_{NI}} \\
&\quad + \frac{5\left[2b_2 N(c_\mathrm{b} - b_{1\mathrm{n}}) - d_2(N+2)^2(b_1 - b_{1\mathrm{n}}) + 2(N+2)b_2 d_1 - 4b_2 b_1\right]^2}{16 b_2 [4b_2 + (N+2)^2 d_2]^2 \bar{x}_{NI}}, \\
f^*_{I_N} &= \frac{[2b_2 N - d_2(N+2)^2](c_\mathrm{b} - b_{1\mathrm{n}}) + 2(N+2)b_2 d_1}{d_2(N+2)^2 + 4b_2}.
\end{aligned}
$$

Firms of type 2 innovate (I) if firm 1 offered N licenses and produce quantities

$$
\begin{aligned}
x^*_{2\mathrm{g}}(NI) &= \frac{c_\mathrm{b} - b_1 + d_1}{2[b_2 + d_2(N+1)]}, \\
x^*_{2\mathrm{g}}(I_N) &= \frac{(N+2)(c_\mathrm{b} - b_{1\mathrm{n}} + d_1)}{2[d_2(N+2)^2 + 4b_2]}.
\end{aligned}
$$

Proof: see Appendix on p.191.

7.2.3 Comparison between subsidy and quota-based policies

In Madlener et al. (2009) it is shown that, in perfectly competitive markets, subsidy and quota policies are equivalent in terms of social welfare maximization. In this study, we have particularly shown that in the subgame-perfect equilibria all fine levels correspond to the subsidy levels.

However, the allocation of welfare to producer vs. consumer surplus differs under these two alternative policies. In particular, the profits achieved by the potential innovator as well as by its competitors are lower under the quota policy (π^q) than under the subsidy policy (π^s) regime:

$$
\begin{aligned}
\pi^q_1(NI) &= \pi^s_1(NI) - f^*(NI)\bar{x}_{NI}; \\
\pi^q_2(NI) &= \pi^s_2(NI) - f^*(NI)\bar{x}_{NI}; \\
\pi^q_1(I_N) &= \pi^s_1(I_N) - f^*(I_N)\bar{x}_{I_N}; \\
\pi^q_2(I_N) &= \pi^s_2(I_N) - f^*(I_N)\bar{x}_{I_N}.
\end{aligned}
$$

Thus, given a no pre-commitment policy, the firms have a strict preference for price rather then quantity controls.

Next, we want to investigate under which policy regime (subsidy or quota) the incentives to innovate are higher. Therefore, we compare the differences of profits of the potential innovator (firm 1) with or without innovation under both regimes. For the subsidy policy, this gain from innovation amounts to

$$\Delta\pi_1^s = \pi_1^s(I_N) - \pi_1^s(NI).$$

Under the quota policy, the corresponding profit difference is

$$\Delta\pi_1^q = \pi_1^q(I_N) - \pi_1^q(NI).$$

The incentives to innovate are higher under the subsidy policy if

$$\Delta\pi_1^s - \Delta\pi_1^q = f_{I_N}^* \bar{x}_{I_N} - f_{NI}^* \bar{x}_{NI} > 0. \tag{7.6}$$

Suppose that the difference between the quota levels $(\bar{x}_{I_N} - \bar{x}_{NI})$ is sufficiently small. Then it can be shown that under the assumption that $d_2(N+1)(c_\mathrm{b} - b_{1\mathrm{n}}) < b_2 d_1$, as made in our model, condition (7.6) is satisfied. Therefore, not only is the subsidy policy preferred by profit-maximizing firms but it also provides a higher incentive to innovate, which is an interesting finding.

7.3 Optimal policy in the presence of innovation: pre-commitment case

In the pre-commitment case, the government is assumed to stick to its green electricity policy (in terms of subsidy and quota) even under innovation. Possible reasons for pre-commitment include: imperfect information, limited ability for short-run policy adjustments etc. Compared to the no pre-commitment assumption, pre-commitment appears to be more realistic, because in the real world there are always difficulties in adjusting policies, for reasons like the ones described above. Besides, there may be other costs associated with policy adjustment, similar to the menu costs in the price adjustment case, that further stymies quick policy reaction to innovations.

We maintain the basic assumptions made in the no pre-commitment case, except that the quota and subsidy levels remain unchanged after the innovation has occurred.

7.3.1 Subsidy policy

We consider an extensive-form game presented in Fig. 7.3. There are two decision stages.

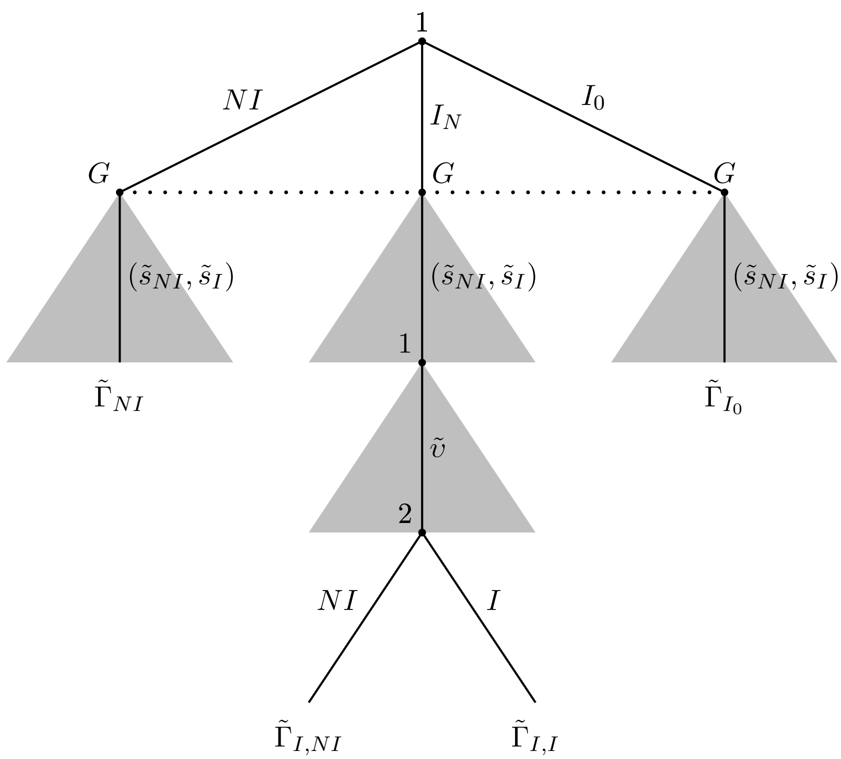


Figure 7.3: Extensive-form game representation, pre-commitment case, subsidy policy

Stage I. Firm 1 decides either not to innovate (NI), to innovate and offer N licenses in the competition stage II (I_N), or to innovate and offer no licenses in stage II (I_0). Simultaneously, the government determines the subsidy levels $\tilde{s}_{NI}$ for non-innovating and $\tilde{s}_I$ for innovating firms in order to maximize social welfare.

Stage II. Given the innovation decision of firm 1 and the decision of the government about the subsidy level, firms 1, 2, ..., $N + 1$ compete in quantities.

(IIa) If firm 1 did not innovate, all firms have identical cost functions $C_{\mathrm{g}}(\cdot)$ and compete in quantities given the subsidy level $\tilde{s}_{NI}$ per unit of green electricity (subgame $\tilde{\Gamma}_{NI}$).

(IIb) If firm 1 did innovate and announced to offer N licenses in stage II, then it first offers licenses to N competitors in return of a royalty $\tilde{v}$ given the subsidy levels $\tilde{s}_I$ and $\tilde{s}_{NI}$. Firms 2, 3, ..., $N+1$ can either accept or reject this offer. Since firms 2, 3, ..., $N+1$ are identical, we assume that either all of them will reject the offer and operate with cost function $C_{\mathrm{g}}(\cdot)$ (competition in quantities will take place in subgame $\tilde{\Gamma}_{I,NI}$) or all of them will accept it and operate with cost function $C_{\mathrm{gn}}(\cdot)$ (competition in subgame $\tilde{\Gamma}_{I,I}$).

(IIc) If firm 1 did innovate but announced that it will offer no licenses in stage 3, then firm 1, operating with cost function $C_{\mathrm{gn}}(\cdot)$, and firms 2, 3, ..., $N+1$, operating with cost function $C_{\mathrm{g}}(\cdot)$, compete in quantities given their subsidy levels $\tilde{s}_I$ and $\tilde{s}_{NI}$, respectively.

Proposition 7.3.1. *There exist two sets of subgame-perfect equilibria in the innovation game with subsidy control and pre-commitment policy. The subgame-perfect equilibrium strategies on the equilibrium path of these two sets are given as follows.*

Set 1. *Firm 1 does not innovate (NI) and produces quantity*

$$x^*_{1\mathrm{g}}(NI,(\tilde{s}^{*1}_{NI},\tilde{s}^{*1}_{I})) = \frac{c_{\mathrm{b}}-b_1+d_1}{2[b_2+d_2(N+1)]}$$

of green electricity. Government chooses subsidy levels $(\tilde{s}^{*1}_{NI},\tilde{s}^{*1}_{I})$ such that

$$\begin{aligned}\tilde{s}^{*1}_{NI} &= \frac{b_2d_1-d_2(N+1)(c_{\mathrm{b}}-b_1)}{b_2+d_2(N+1)};\\ \tilde{s}^{*1}_{I} &\in \left[-(c_{\mathrm{b}}-b_{1\mathrm{n}})\pm\sqrt{\frac{2b_2}{N+2}\left(\frac{b_2(c_{\mathrm{b}}-b_1+d_1)^2}{[b_2+d_2(N+1)]^2}+4r(\Delta b_1)^2\right)}\,\right].\end{aligned}$$

Firms of type 2 produce quantity

$$x^*_{2\mathrm{g}}(NI,(\tilde{s}^{*1}_{NI},\tilde{s}^{*1}_{I})) = \frac{c_{\mathrm{b}}-b_1+d_1}{2[b_2+d_2(N+1)]}$$

of green electricity.

Set 2. *Firm 1 innovates and offers N licenses (I_N) in return of a royalty*

$$\tilde{v}^* = \frac{b_2(N+2)(c_{\mathrm{b}} - b_1 + d_1)}{d_2(N+2)^2 + 4b_2}$$

per unit of green electricity produced by firms of type 2 and itself produces quantity

$$x^*_{1\mathrm{g}}(I_N, (\tilde{s}^{*2}_{NI}, \tilde{s}^{*2}_{I}) = \frac{(N+2)(c_{\mathrm{b}} - b_{1\mathrm{n}} + d_1)}{d_2(N+2)^2 + 4b_2}$$

of green electricity. Government sets subsidy levels ($\tilde{s}^{*2}_{NI}, \tilde{s}^{*2}_{I}$) such that

$$\begin{aligned}
\tilde{s}^{*2}_{NI} &\in \left[-(c_{\mathrm{b}} - b_1) \pm \sqrt{b_2\left(\frac{b_2(N+2)^3(c_{\mathrm{b}} - b_{1\mathrm{n}} + d_1)^2}{[4b_2 + d_2(N+2)^2]^2} - 4r(\Delta b_1)^2\right)}\right]; \\
\tilde{s}^{*2}_{I} &= \frac{[2b_2N - d_2(N+2)^2](c_{\mathrm{b}} - b_{1\mathrm{n}}) + 2(N+2)b_2d_1}{d_2(N+2)^2 + 4b_2}.
\end{aligned}$$

Firms of type 2 innovate (I) and produce quantity

$$x^*_{2\mathrm{g}}(I_N, (\tilde{s}^{*2}_{NI}, \tilde{s}^{*2}_{I}) = \frac{(N+2)(c_{\mathrm{b}} - b_{1\mathrm{n}} + d_1)}{2[d_2(N+2)^2 + 4b_2]}$$

of green electricity.

Proof: see Appendix on p.191.

7.3.2 Quota-based policy

Now we consider an extensive-form game with the structure presented in Figure 7.4. As under the subsidy policy, there are two decision stages.

Stage I. Firm 1 decides either not to innovate (NI), to innovate and offer N licenses in the competition stage II (I_N), or to innovate and offer no licenses in stage II (I_0). Simultaneously, the government determines the fine levels $\tilde{f}_{NI}$ for non-innovating and $\tilde{f}_I$ for innovating firms in order to maximize social welfare.

Stage II. Given the innovation decision of firm 1 and the decision of the government about the subsidy level, firms 1, 2, ..., $N+1$ compete in quantities.

(IIa) If firm 1 did not innovate, all firms have identical cost functions $C_{\mathrm{g}}(\cdot)$ and compete in quantities given the fine level $\tilde{f}_{NI}$ per unit of green electricity (subgame $\tilde{\Gamma}^q_{NI}$).

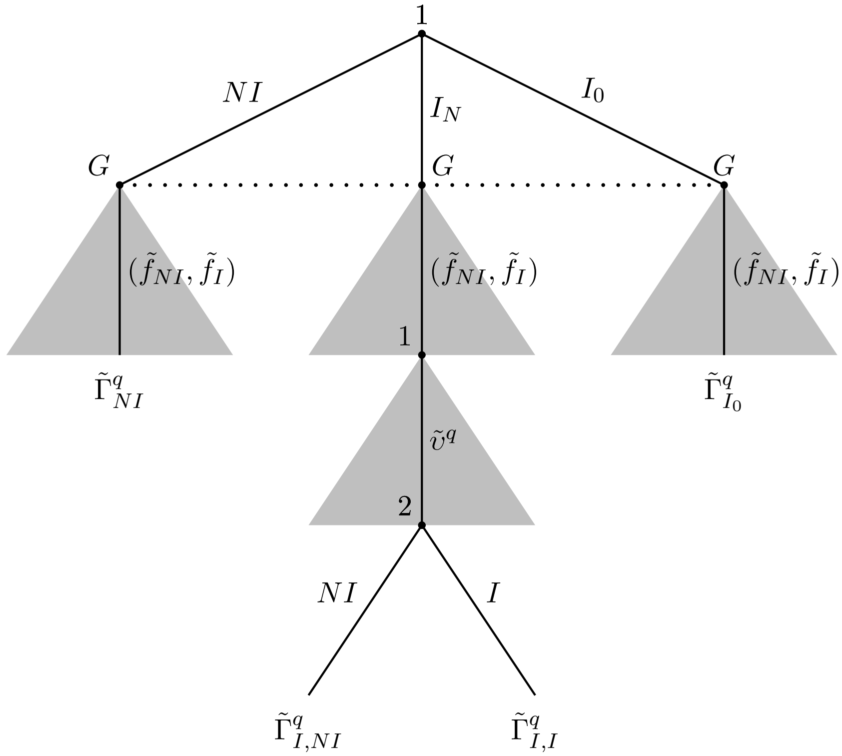


Figure 7.4: Extensive-form game representation, pre-commitment case, quota policy

(IIb) If firm 1 did innovate and announced to offer N licenses in stage II, then it first offers licenses to N competitors in return of a royalty $\tilde{v}^q$ given the fine levels $\tilde{f}_I$ and $\tilde{f}_{NI}$. Firms 2, 3, …, $N+1$ can either accept or reject this offer. Since firms 2, 3, …, $N+1$ are identical, we assume that either all of them will reject the offer and operate with cost function $C_{\mathrm{g}}(\cdot)$ (competition in quantities will take place in subgame $\tilde{\Gamma}^q_{I,NI}$) or all of them will accept it and operate with cost function $C_{\mathrm{gn}}(\cdot)$ (competition in subgame $\tilde{\Gamma}^q_{I,I}$).

(IIc) If firm 1 did innovate but announced that it will offer no licenses in stage 3, then firm 1, operating with cost function $C_{\mathrm{gn}}(\cdot)$, and firms 2, 3, …, $N+1$, operating with cost function $C_{\mathrm{g}}(\cdot)$, compete in quantities given their fine levels $\tilde{f}_I$ and $\tilde{f}_{NI}$, respectively.

Proposition 7.3.2. *There exist two sets of subgame-perfect equilibria strategies in the pre-commitment game with quotas. The subgame-perfect equilibrium strategies on the equilibrium path of this game are given as follows.*

Set 1. *Firm 1 does not innovate (NI) and produces quantity*

$$x^*_{1\mathrm{g}}(NI, (\tilde{f}^{*1}_{NI}, \tilde{f}^{*1}_{I})) = \frac{c_\mathrm{b} - b_1 + d_1}{2[b_2 + d_2(N+1)]}$$

of green electricity. Government sets fine levels ($\tilde{f}^{*1}_{NI}, \tilde{f}^{*1}_{I}$) such that

$$\begin{aligned}
\tilde{f}^{*1}_{NI} &= \frac{b_2 d_1 - d_2(N+1)(c_\mathrm{b} - b_1)}{b_2 + d_2(N+1)};\\
\tilde{f}^{*1}_{I} &\in \left[-(c_\mathrm{b} - b_{1\mathrm{n}}) + 4b_2\bar{x}_I \pm \frac{\xi\sqrt{2b_2}}{(N+2)[b_2 + (N+1)d_2]}\right],
\end{aligned}$$

where $\bar{x}_I$ denotes the minimum quota to be produced by an innovating firm, and

$$\xi = \sqrt{8b_2^3\bar{x}_I^2 + 4(N+1)^2(N+2)d_2^2\delta + 4b_2^2\epsilon + b_2\eta},$$

with

$$\begin{aligned}
\delta &= r(\Delta b_1)^2 - (c_\mathrm{b} - b_{1\mathrm{n}})\bar{x}_I - (c_\mathrm{b} - b_1)\bar{x}_{NI};\\
\epsilon &= (N+2)r(\Delta b_1)^2 + 4(N+1)d_2\bar{x}_I^2 - (N+2)[(c_\mathrm{b} - b_{1\mathrm{n}})\bar{x}_I + d_1\bar{x}_{NI}];\\
\eta &= (N+2)[(c_\mathrm{b} - b_1 + d_1)^2 - 2c_\mathrm{b}d_1] +\\
&\quad +4(N+1)d_2[2(N+2)r(\Delta b_1)^2 + 2(N+1)d_2\bar{x}_I^2 + 2b_{1\mathrm{n}}\bar{x}_I - (N+2)(c_\mathrm{b} + b_1)\bar{x}_{NI})].
\end{aligned}$$

Firms of type 2 produce quantity

$$x^*_{2\mathrm{g}}(NI, (\tilde{f}^{*1}_{NI}, \tilde{f}^{*1}_{I})) = \frac{c_\mathrm{b} - b_1 + d_1}{2[b_2 + d_2(N+1)]}$$

of green electricity.

Set 2. *Firm 1 innovates and offers N licenses (I_N) in return of a royalty*

$$\tilde{\upsilon}^{*q} = \frac{b_2(N+2)(c_\mathrm{b} - b_{1\mathrm{n}} + d_1)}{d_2(N+2)^2 + 4b_2}$$

per unit of green electricity produced by firms of type 2 and itself produces quantity

$$x^*_{1\mathrm{g}}(I_N, (\tilde{f}^{*2}_{NI}, \tilde{f}^{*2}_{I}) = \frac{(N+2)(c_\mathrm{b} - b_{1\mathrm{n}} + d_1)}{d_2(N+2)^2 + 4b_2}$$

of green electricity. Government sets fine levels $(\tilde{f}^{*2}_{NI}, \tilde{f}^{*2}_{I})$ *such that*

$$\begin{aligned}
\tilde{f}^{2}_{NI} &\in \left[-(c_{\rm b}-b_{1\rm n})+2b_2\bar{x}_{NI} \pm \frac{\psi\sqrt{2b_2}}{4b_2+(N+2)^2d_2}\right];\\
\tilde{f}^{*2}_{I} &= \frac{[2b_2N-d_2(N+2)^2](c_{\rm b}-b_{1\rm n})+2(N+2)b_2d_1}{d_2(N+2)^2+4b_2},
\end{aligned}$$

where $\bar{x}_{NI}$ *denotes the minimum quota to be produced by a non-innovating firm, and*

$$\psi = \sqrt{32b_2^3\bar{x}^2_{NI} - 2(N+2)^4\kappa - 16b_2^2\mu + (N+2)^2b_2\nu + 2b_{1\rm n}\rho + 2d_2\tau}$$

with

$$\begin{aligned}
\kappa &= r(\Delta b_1)^2 - (c_{\rm b}-b_{1\rm n})\bar{x}_I - (c_{\rm b}-b_1)\bar{x}_{NI};\\
\mu &= 2r(\Delta b_1)^2 + [N(c_{\rm b}-b_{1\rm n}+d_1)+2d_1]\bar{x}_I + [2(c_{\rm b}-b_1)-(N+2)^2d_2\bar{x}_{NI}]\bar{x}_{NI};\\
\nu &= (N+2)[(c_{\rm b}-b_{1\rm n}+d_1)^2 + 2b_{1\rm n}(c_{\rm b}+d_1) - 4(N+2)d_2\bar{x}_I;\\
\rho &= 2(N-2)d_2\bar{x}_I - (N+2)(c_{\rm b}+d_1);\\
\tau &= \bar{x}_{NI}[(N+2)^2d_2\bar{x}_{NI} - 8(c_{\rm b}-b_1)] - 2c_{\rm b}(N-2)\bar{x}_I - 8r(\Delta b_1)^2.
\end{aligned}$$

Firms of type 2 innovate (I) and each produce quantity

$$x^*_{2\rm g}(I_N, (\tilde{f}^{*2}_{NI}, \tilde{f}^{*2}_{I}) = \frac{(N+2)(c_{\rm b}-b_{1\rm n}+d_1)}{2[d_2(N+2)^2+4b_2]}$$

of green electricity.

Proof: see Appendix on p.193.

7.3.3 Comparison between subsidy and quota-based policy

Under pre-commitment, the subsidy and quota policies again are equivalent in terms of social welfare. However, the firms prefer the subsidy policy since they achieve higher profits than under the quota policy.

Furthermore, as under no pre-commitment, the profits achieved by the potential innovator as well as by its competitors are lower under the quota policy (π^q) than under the subsidy policy (π^s):

$$\begin{aligned}
\pi_1^q(NI) &= \pi_1^s(NI) - \tilde{f}^*(NI)\bar{x}_{NI}; \\
\pi_2^q(NI) &= \pi_2^s(NI) - \tilde{f}^*(NI)\bar{x}_{NI}; \\
\pi_1^q(I_N) &= \pi_1^s(I_N) - \tilde{f}^*(I_N)\bar{x}_{I_N}; \\
\pi_2^q(I_N) &= \pi_2^s(I_N) - \tilde{f}^*(I_N)\bar{x}_{I_N}.
\end{aligned}$$

Thus, the firms have a strict preference for the subsidy policy under pre-commitment, too.

Again, as in the no pre-commitment case, the innovation incentives are higher under the subsidy policy:

$$\Delta\pi_1^s - \Delta\pi_1^q = \tilde{f}_{I_N}^*\bar{x}_{I_N} - \tilde{f}_{NI}^*\bar{x}_{NI} > 0. \tag{7.7}$$

7.4 Discussion and Conclusions

Madlener et al. (2009) found that the conventional wisdom related to the equivalence of tax (subsidy) and quota (certificate) schemes in terms of static efficiency may not hold if markets for electric power are imperfectly competitive. Due to the inequivalence found in terms of social welfare, the authors recommend targeted subsidies as being the preferable policy instrument.

In this paper, we have followed up studying the merits of price and quantity control policies for promoting renewable electricity generation. In particular, we study the role of government regulatory pre-commitment when technical innovation is present.

In the pre-commitment case, the government is assumed to stick to its green electricity policy (in terms of subsidy and quota) even under innovation. Possible reasons for pre-commitment include: imperfect information, limited ability for short-run policy adjustments etc. Compared to the no pre-commitment assumption, pre-commitment appears to be more realistic, because in the real world there are always difficulties in adjusting policies, for reasons like the ones described above. Besides, there may be other costs associated with policy adjustment, similar to the menu costs in the price adjustment case, that further stymies quick policy reaction to innovations.

We maintain the basic assumptions made in the no pre-commitment case, except that the quota and subsidy levels remain unchanged after the innovation occurred. We can conclude that the difference is larger without pre-commitment, i.e. the subsidy scheme is preferred more in the case of no pre-commitment.

Thus we find that the price (subsidy) policy is again preferred in terms of promoting innovation of green electricity technology. The intuition behind the result is also the same as that under the no pre-commitment case. Since technological improvement and innovation mainly represent the dynamic aspect of energy efficiency for a firm (and also for an economy), our results strongly support the subsidy policy in terms of its dynamic efficiency in general, no matter which policy regime, pre-commitment or no pre-commitment, is feasible (or followed) in the real world.

An important finding concerns the issue whether the existence of equilibrium solutions depend on pre-commitment. The sets of subgame-perfect equilibria derived in this paper confirm that pre-commitment can influence the equilibrium conditions. In particular, under no pre-commitment a sufficiently high cost reduction would necessarily lead to innovation and exclude the possibility that no innovation occurs. By way of contrast, both equilibria are possible under pre-commitment even if the cost reduction by the innovation is high. Still, under pre-commitment an equilibrium with innovation remains possible in a case of a relatively low cost reduction as opposed to the no pre-commitment case. It follows that a government with a preference for innovations being performed if the achievable cost reduction is high (and otherwise not) should be in favor of the no pre-commitment regime.

Acknowledgements

The authors gratefully acknowledge comments received on earlier versions of this manuscript from Peter Zweifel, Eberhard Feess, Michael Kuenzle and Julia Meyer, as well as participants of the SURED 2006 conference in Ascona, Switzerland, 5–9 June 2006.

7.5 Appendix

Proof of Lemma 7.2.1 Suppose that firm 1 does not innovate in stage I, i.e. it chooses action NI. The subsidy level chosen by the government for all firms in stage II is s_{NI}. Given a competitive market in stage III, a representative power generator i faces the optimization problem

$$\max_{x_{i\mathrm{b}},x_{i\mathrm{g}}} [px_{i\mathrm{b}} + (p + s_{NI})x_{i\mathrm{g}} - C_{\mathrm{b}}(x_{i\mathrm{b}}) - C_{\mathrm{g}}(x_{i\mathrm{g}})], \tag{7.8}$$

where $x_{i\mathrm{b}}$ and $x_{i\mathrm{g}}$ denote the amounts of electricity produced by firm i from fossil/nuclear ('brown') and renewable ('green') energy sources, respectively, and p, the average market price for electricity. The f.o.c. for an interior solution are

$$p \quad -C'_{\mathrm{b}}[x^*_{i\mathrm{b}}] = 0 \tag{7.9}$$

$$p \quad +s_{NI} - C'_{\mathrm{g}}[x^*_{i\mathrm{g}}] = 0. \tag{7.10}$$

Inserting (7.9) into (7.10) reveals that in an optimum with $x_{i\mathrm{b}} > 0$ and $x_{i\mathrm{g}} > 0$, the government subsidy s_{NI} has to be equal to the difference (in absolute terms) between $C'_{\mathrm{g}}[x^*_{i\mathrm{g}}]$ and c_{b}, i.e. the marginal costs of green electricity evaluated at the optimum and the constant marginal cost of brown electricity. The intuition behind this result from an economic perspective is that if $s_{NI} > C'_{\mathrm{g}}[x^*_{i\mathrm{g}}] - c_{\mathrm{b}}$, then all generators will exclusively supply green electricity. In contrast, if $s_{NI} < C'_{\mathrm{g}}[x^*_{i\mathrm{g}}] - c_{\mathrm{b}}$, no green electricity at all will be provided. Given the assumptions of a competitive market and homogeneous costs, the subgame solution is described by (7.9) and (7.10). In particular, all firms produce the same quantity of green electricity given by

$$x_{i\mathrm{g}}(NI, s_{NI}) \quad = \quad \frac{c_{\mathrm{b}} - b_1 + s_{NI}}{2b_2}, \tag{7.11}$$

while each firm's profit amounts to

$$\pi_i(NI, s_{NI}) \quad = \quad \frac{(c_{\mathrm{b}} - b_1 + s_{NI})^2}{4b_2}. \tag{7.12}$$

Proof of Lemma 7.2.2 Suppose that firm 1 innovates and announces to offer N licenses. The government determines welfare-maximizing subsidy levels s_{NI_N} for non-innovating and s_{I_N} for innovating firms. We denote the royalty for the new technology per unit of green power as υ. In equilibrium, it must not exceed the cost difference $C_{\mathrm{g}}(x_{2\mathrm{g}}) - C_{\mathrm{gn}}(x_{2\mathrm{g}})$, as otherwise there is no incentive to switch to the new technology.

Subgame $\Gamma_{I_N,NI}$. Suppose that firms 2, 3, …, $N+1$ (from here on: firms of type 2) rejected firm 1's offer. Then firm 1 operates with the new cost function $C_{\mathrm{gn}}(x_{1\mathrm{g}})$ while firms of type 2 continue to operate with the cost function $C_{\mathrm{g}}(x_{2\mathrm{g}})$. Thus, firm 1's profit maximization problem is given by

$$\max_{x_{1\mathrm{b}},x_{1\mathrm{g}}} [px_{1\mathrm{b}} + (p + s_{I_N})x_{1\mathrm{g}} - C_{\mathrm{b}}(x_{1\mathrm{b}}) - C_{\mathrm{gn}}(x_{1\mathrm{g}}) - R(b_1 - b_{1\mathrm{n}})], \tag{7.13}$$

while firm 2's profit maximization problem corresponds to (7.8) with $i = 2$ and $s_{NI} = s_{NI_N}$. Thus, quantities of green electricity produced by firm 1 and firms of type 2 are given by

$$\begin{aligned} x_{1\mathrm{g}}(I_N, (s_{NI_N}, s_{I_N}), \upsilon, NI) &= \frac{c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_N}}{2b_2}; \\ x_{2\mathrm{g}}(I_N, (s_{NI_N}, s_{I_N}), \upsilon, NI) &= \frac{c_\mathrm{b} - b_1 + s_{NI_N}}{2b_2}, \end{aligned}$$

and firms' profits therefore amount to

$$\begin{aligned} \pi_1(I_N, (s_{NI_N}, s_{I_N}), \upsilon, NI) &= \frac{(c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_N})^2}{4b_2} - R(b_1 - b_{1\mathrm{n}}); \\ \pi_2(I_N, (s_{NI_N}, s_{I_N}), \upsilon, NI) &= \frac{(c_\mathrm{b} - b_1 + s_{NI_N})^2}{4b_2}. \end{aligned}$$

Subgame $\Gamma_{I_N,I}$. Now suppose that firms of type 2 accept firm 1's offer and pay a royalty of υ per unit of green electricity produced. Then all firms operate with the new cost function $C_{\mathrm{gn}}(x_\mathrm{g})$. The profit maximization problems of firm 1 and firms of type 2 are respectively given by

$$\max_{x_{1\mathrm{b}}, x_{1\mathrm{g}}} [px_{1\mathrm{b}} + (p + s_{I_N})x_{1\mathrm{g}} - C_\mathrm{b}(x_{1\mathrm{b}}) - C_{\mathrm{gn}}(x_{1\mathrm{g}}) + N\upsilon x_{2\mathrm{g}} - R(b_1 - b_{1\mathrm{n}})]; \tag{7.14}$$

$$\max_{x_{2\mathrm{b}}, x_{2\mathrm{g}}} [px_{2\mathrm{b}} + (p + s_{I_N})x_{2\mathrm{g}} - C_\mathrm{b}(x_{2\mathrm{b}}) - C_{\mathrm{gn}}(x_{2\mathrm{g}}) - \upsilon x_{2\mathrm{g}}]; \tag{7.15}$$

the quantities of green electricity produced are

$$\begin{aligned} x_{1\mathrm{g}}(I_N, (s_{NI_N}, s_{I_N}), \upsilon, I) &= \frac{c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_N}}{2b_2}; \\ x_{2\mathrm{g}}(I_N, (s_{NI_N}, s_{I_N}), \upsilon, I) &= \frac{c_\mathrm{b} - (b_{1\mathrm{n}} + \upsilon) + s_{I_N}}{2b_2}. \end{aligned}$$

The firms' profits thus amount to

$$\begin{aligned} \pi_1(I_N, (s_{NI_N}, s_{I_N}), \upsilon, I) &= \frac{(c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_N})^2}{4b_2} + N\upsilon\frac{c_\mathrm{b} - (b_{1\mathrm{n}} + \upsilon) + s_{I_N}}{2b_2} - R(b_1 - b_{1\mathrm{n}}); \\ \pi_2(I_N, (s_{NI_N}, s_{I_N}), \upsilon, I) &= \frac{(c_\mathrm{b} - (b_{1\mathrm{n}} + \upsilon) + s_{I_N})^2}{4b_2}. \end{aligned}$$

Firms of type 2 decide in stage IIIb whether to reject (NI) or accept (I) the offer, depending on the comparison of the maximum profits calculated for subgames $\Gamma_{I_N,NI}$ and $\Gamma_{I_N,I}$. Thus, their subgame-perfect equilibrium actions are given as follows:

$$\begin{cases} NI & \text{if} \quad \upsilon > \upsilon^{\max} := (b_1 - b_{1\mathrm{n}}) - (s_{NI_N} - s_{I_N}) \\ I & \text{if} \quad \text{otherwise.} \end{cases}$$

In other words, $\upsilon^{\max}$ is the highest possible royalty level at which firms of type 2 innovate.

Firm 1's decision in stage IIIb is based on the maximization of its profits w.r.t. royalty level υ. Notice that firm 1's profit, provided firms of type 2 accept the offer $\pi_1(I_N, (s_{NI_N}, s_{I_N}), \upsilon, I)$, is always at least as high as if they reject it as long as $\upsilon \in [0, c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_N}]$. Moreover, the profit function $\pi_1(I_N, (s_{NI_N}, s_{I_N}), \upsilon, I)$ attains its maximum in υ at the royalty level $\upsilon = (c_{\rm b} - b_{1\rm n} + s_{I_N})/2$. Thus, taking into consideration the possible case of a corner solution, firm 1's equilibrium offer υ^* in stage IIIb is given by

$$\upsilon^* = \begin{cases} \upsilon^{\max} = (b_1 - b_{1\rm n}) - (s_{NI_N} - s_{I_N}) & \text{if } (b_1 - b_{1\rm n}) - (s_{NI_N} - s_{I_N}) < \dfrac{c_{\rm b} - b_{1\rm n} + s_{I_N}}{2}; \\ \dfrac{c_{\rm b} - b_{1\rm n} + s_{I_N}}{2} & \text{otherwise.} \end{cases}$$

This offer will always be accepted by a firm of type 2 in the equilibrium[9]. Green electricity produced by firm 2 in the subgame starting at node G_2 thus amounts to

$$x_{2\rm g}(I_N, (s_{NI_N}, s_{I_N}), I) = \begin{cases} \dfrac{(c_b - b_1 + s_{NI_N})}{2b_2} & \text{if } (b_1 - b_{1\rm n}) - (s_{NI_N} - s_{I_N}) < \dfrac{c_{\rm b} - b_{1\rm n} + s_{I_N}}{2}; \\ \dfrac{c_{\rm b} - b_{1\rm n} + s_{I_N}}{4b_2} & \text{otherwise.} \end{cases}$$

Firms' profits in this subgame are thus given by

$$\pi_1(I_N, (s_{NI_N}, s_{I_N}), I) \quad = \tag{7.16}$$

$$= \begin{cases} \dfrac{(c_b - b_{1\rm n} + s_{I_N})^2}{4b_2} - R(b_1 - b_{1\rm n}) + \\ +N[(b_1 - b_{1\rm n}) - (s_{NI_N} - s_{I_N})]\frac{c_b - b_1 + s_{NI_N}}{2b_2} & \text{if } (b_1 - b_{1\rm n}) - (s_{NI_N} - s_{I_N}) < \dfrac{c_{\rm b} - b_{1\rm n} + s_{I_N}}{2}; \\ \dfrac{(N+2)(c_{\rm b} - b_{1\rm n} + s_{I_N})^2}{8b_2} - R(b_1 - b_{1\rm n}) & \text{otherwise,} \end{cases}$$

$$\pi_2(I_N, (s_{NI_N}, s_{I_N}), I) = \begin{cases} \dfrac{(c_b - b_1 + s_{NI_N})^2}{4b_2} & \text{if } (b_1 - b_{1\rm n}) - (s_{NI_N} - s_{I_N}) < \dfrac{c_{\rm b} - b_{1\rm n} + s_{I_N}}{2}; \\ \dfrac{(c_{\rm b} - b_{1\rm n} + s_{I_N})^2}{8b_2} & \text{otherwise.} \end{cases}$$

Proof of Lemma 7.2.3 Suppose that firm 1 innovates but offers no licenses to competitors (I_0). The government determines welfare-maximizing subsidy levels (s_{NI_0}, s_{I_0}). Firm 1, after innovating, operates with the new cost function $C_{\rm gn}(x_{1\rm g})$ and firms of type 2 continue to operate with the cost function $C_{\rm g}(x_{2\rm g})$. Thus, firm 1's profit maximization problem is given by

$$\max_{x_{1\rm b}, x_{1\rm g}} [px_{1\rm b} + (p + s_{I_0})x_{1\rm g} - C_{\rm b}(x_{1\rm b}) - C_{\rm gn}(x_{1\rm g}) - R(b_1 - b_{1\rm n})], \tag{7.17}$$

while firm 2's profit maximization problem corresponds to (7.8) with $i = 2$ and $s_{NI} = s_{NI_0}$. The quantities of green electricity produced by firm 1 and firms of type 2 are therefore given by

$$\begin{aligned} x_{1\rm g}(I_0, (s_{NI_0}, s_{I_0})) &= \frac{c_{\rm b} - b_{1\rm n} + s_{I_0}}{2b_2}; \\ x_{2\rm g}(I_0, (s_{NI_0}, s_{I_0})) &= \frac{c_{\rm b} - b_1 + s_{NI_0}}{2b_2}. \end{aligned}$$

[9] As usual, we assume that in a case of indifference firms of type 2 decide in favor of the adoption of the new technology.

Firms' profits therefore amount to

$$\begin{aligned}
\pi_1(I_0,(s_{NI_0},s_{I_0})) &= \frac{(c_b - b_{1n} + s_{I_0})^2}{4b_2} - R(b_1 - b_{1n}); &(7.18)\\
\pi_2(I_0,(s_{NI_0},s_{I_0})) &= \frac{(c_b - b_1 + s_{NI_0})^2}{4b_2}. &(7.19)
\end{aligned}$$

Proof of Lemma 7.2.4 Given the decision of firm 1 not to innovate, the government anticipates all firms' optimal quantity decisions in the subgame Γ_{NI} and maximizes the social welfare function

$$\begin{aligned}
W_{NI}(s_{NI}) &= Q\left(a - \frac{Q}{2}\right) + (N+1)\pi_i(NI,s_{NI}) - s_{NI}(N+1)x_g(NI,s_{NI})\\
&\quad + d_1(N+1)x_g(NI,s_{NI}) - d_2[(N+1)x_g(NI,s_{NI})]^2 &(7.20)\\
&= \frac{(a-c_b)(a+c_b)}{2} + (N+1)\,\frac{(c_b - b_1 + s_{NI})^2}{4b_2} - (N+1)s_{NI}\frac{c_b - b_1 + s_{NI}}{2b_2}\\
&\quad + (N+1)d_1\,\frac{c_b - b_1 + s_{NI}}{2b_2} - d_2(N+1)^2\frac{(c_b - b_1 + s_{NI})^2}{4b_2^2} &(7.21)
\end{aligned}$$

with respect to s_{NI}. The socially optimal subsidy level is thus given by

$$s^*_{NI} = \frac{b_2 d_1 - d_2(N+1)(c_b - b_1)}{b_2 + d_2(N+1)}, \qquad (7.22)$$

while the equilibrium quantities and profits are

$$\begin{aligned}
x_{ig}(NI) &= \frac{c_b - b_1 + d_1}{2[b_2 + d_2(N+1)]};\\
\pi_i(NI) &= \frac{b_2(c_b - b_1 + d_1)^2}{4[b_2 + d_2(N+1)]^2}.
\end{aligned}$$

Proof of Lemma 7.2.5 Given the decision of firm 1 to innovate and offer N licenses, the government anticipates the equilibrium outcomes of subgames $\Gamma_{I_N,NI}$, $\Gamma_{I_N,I}$, as well as that of the royalty bargaining subgame, and faces the social welfare function

$$\begin{aligned}
W_{I_N}(s_{NI_N},s_{I_N}) &= Q\left(a - \frac{Q}{2}\right) + \pi_1(I_N,(s_{NI_N},s_{I_N})) + N\pi_2(I_N,(s_{NI_N},s_{I_N}))\\
&\quad - s_{I_N}x_{1g}(NI,(s_{NI_N},s_{I_N})) - Ns_{I_N}x_{2g}(NI,(s_{NI_N},s_{I_N}))\\
&\quad + d_1\left[x_{1g}(NI,(s_{NI_N},s_{I_N})) + Nx_{2g}(NI,(s_{NI_N},s_{I_N}))\right]\\
&\quad - d_2\left[x_{1g}(NI,(s_{NI_N},s_{I_N})) + Nx_{2g}(NI,(s_{NI_N},s_{I_N}))\right]^2.
\end{aligned}$$

Since the outcome of the following subgame crucially depends on whether condition

$$(b_1 - b_{1n}) - (s_{NI_N} - s_{I_N}) \geq \frac{c_b - b_{1n} + s_{I_N}}{2} \qquad (7.23)$$

is satisfied, the welfare function in stage IIb is a piecewise-defined continuous function. We distinguish two cases, depending on whether or not condition (7.23) is fulfilled.

Case 1: condition (7.23) is satisfied. The government maximizes the welfare function

$$\begin{aligned} W_{I_N}(s_{NI_N}, s_{I_N}) &= \frac{(a-c_\mathrm{b})(a+c_\mathrm{b})}{2} + 2(N+1)\,\frac{(c_\mathrm{b}-b_{1\mathrm{n}}+s_{I_N})^2}{8b_2} - R(b_1-b_{1\mathrm{n}}) \\ &\quad - s_{I_N}(N+2)\frac{c_\mathrm{b}-b_{1\mathrm{n}}+s_{I_N}}{4b_2} \\ &\quad + d_1(N+2)\frac{c_\mathrm{b}-b_{1\mathrm{n}}+s_{I_N}}{4b_2} - d_2(N+2)^2\,\frac{(c_\mathrm{b}-b_{1\mathrm{n}}+s_{I_N})^2}{16b_2^2} \end{aligned}$$

with respect to (s_{NI_N}, s_{I_N}) and subject to constraint (7.23). The socially optimal subsidy level is given by

$$s^*_{I_N} = \frac{[2b_2N - d_2(N+2)^2](c_\mathrm{b}-b_{1\mathrm{n}}) + 2(N+2)b_2d_1}{d_2(N+2)^2+4b_2}. \tag{7.24}$$

The maximum welfare level attained in this case is

$$W^*(I_N) = \frac{(N+2)^2(c_\mathrm{b}-b_{1\mathrm{n}}+d_1)^2}{4[(N+2)^2d_2+4b_2]} - R(b_1-b_{1\mathrm{n}}).$$

Case 2: condition (7.23) is not satisfied. The government maximizes the welfare function

$$\begin{aligned} W'_{I_N}(s'_{NI_N}, s'_{I_N}) &= \frac{(a-c_\mathrm{b})(a+c_\mathrm{b})}{2} + \frac{(c_\mathrm{b}-b_{1\mathrm{n}}+s'_{I_N})^2}{4b_2} - R(b_1-b_{1\mathrm{n}}) \\ &\quad + N[(b_1-b_{1\mathrm{n}}) - (s'_{NI_N}-s'_{I_N})]\frac{c_\mathrm{b}-b_1+s'_{NI_N}}{2b_2} + N\,\frac{(c_\mathrm{b}-b_1+s'_{NI_N})^2}{4b_2} \\ &\quad - s'_{I_N}\frac{c_\mathrm{b}-b_{1\mathrm{n}}+s'_{I_N}}{2b_2} - Ns'_{I_N}\frac{c_\mathrm{b}-b_1+s'_{NI_N}}{2b_2} \\ &\quad + d_1\left(\frac{c_\mathrm{b}-b_{1\mathrm{n}}+s'_{I_N}}{2b_2} + N\,\frac{c_\mathrm{b}-b_1+s'_{NI_N}}{2b_2}\right) \\ &\quad - d_2\left(\frac{c_\mathrm{b}-b_{1\mathrm{n}}+s'_{I_N}}{2b_2} + N\,\frac{c_\mathrm{b}-b_1+s'_{NI_N}}{2b_2}\right)^2 \end{aligned}$$

with respect to (s'_{NI_N}, s'_{I_N}) and subject to constraint (7.23) reversed with $<$. The socially optimal subsidy levels are given by

$$s'^*_{NI_N} = \frac{b_2d_1 - d_2(N+1)(c_\mathrm{b}-b_1) + b_2(b_1-b_{1\mathrm{n}})}{b_2+d_2(N+1)}; \tag{7.25}$$

$$s'^*_{I_N} = \frac{b_2d_1 - d_2(N+1)(c_\mathrm{b}-b_{1\mathrm{n}})}{b_2+d_2(N+1)}. \tag{7.26}$$

The maximum welfare level to be attained is

$$W'^*(I_N) = \frac{(N+1)(c_\mathrm{b}-b_{1\mathrm{n}}+d_1)^2}{4[b_2+d_2(N+1)]} - R(b_1-b_{1\mathrm{n}}).$$

A simple computation shows that $W^*(I_N) > W'^*(I_N)$ for any $N > 0$. Thus, the optimal decision of the government in stage IIb is given by any combination of subsidies

$$(s^*_{NI_N}, s^*_{I_N}) = \left(s^*_{NI_N}, \frac{[2b_2N - d_2(N+2)^2](c_\mathrm{b}-b_{1\mathrm{n}}) + 2(N+2)b_2d_1}{d_2(N+2)^2+4b_2}\right),$$

where

$$s^*_{NI_N} \geq (b_1 - b_{1\mathrm{n}}) + \frac{[b_2(N-2) - d_2(N+2)^2](c_\mathrm{b} - b_{1\mathrm{n}}) + (N+2)b_2 d_1}{d_2(N+2)^2 + 4b_2}.$$

Proof of Lemma 7.2.6 Given the decision of firm 1 to innovate and offer no licenses, the government anticipates the equilibrium outcome of subgame Γ_{I_0} and maximizes the welfare function

$$\begin{aligned} W_{I_0}(s_{NI_0}, s_{I_0}) \quad = \quad & \frac{(a-c_\mathrm{b})(a+c_\mathrm{b})}{2} + \frac{(c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_0})^2}{4b_2} - R(b_1 - b_{1\mathrm{n}}) + \\ & + N\,\frac{(c_\mathrm{b} - b_1 + s_{NI_0})^2}{4b_2} - s_{I_0}\frac{c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_0}}{2b_2} - N s_{NI_0}\frac{c_\mathrm{b} - b_1 + s_{NI_0}}{2b_2} \\ & + d_1\left(\frac{c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_0}}{2b_2} + N\,\frac{c_\mathrm{b} - b_1 + s_{NI_0}}{2b_2}\right) - \\ & - d_2\left(\frac{c_\mathrm{b} - b_{1\mathrm{n}} + s_{I_0}}{2b_2} + N\,\frac{c_\mathrm{b} - b_1 + s_{NI_0}}{2b_2}\right)^2 \end{aligned}$$

with respect to (s_{NI_0}, s_{I_0}). The socially optimal subsidy levels in this subgame coincide for the innovating firm and the non-innovating firms and are given by

$$s^*_{NI_0} = s^*_{I_0} \quad = \quad \frac{b_2 d_1 - d_2(N+1)(c_\mathrm{b} - b_1) - d_2(b_1 - b_{1\mathrm{n}})}{b_2 + d_2(N+1)}. \tag{7.27}$$

Proof of Proposition 7.2.7 In stage I, firm 1 anticipates optimal decisions of the government and other firms in the subsequent subgames and thus decides whether or not to innovate (and if so, whether to offer licenses) based on its maximum profits to be attained given the utility-maximizing decisions of other players. First of all, observe that, for any $N > 0$, $\pi_1(I_N, (s_{NI_N, s_{I_N}})) > \pi_1(I_0, (s_{NI_0}, s_{I_0}))$. Thus, firm 1 will never take the strictly dominated action I_0 in stage I. Consequently, the solution depends on the comparison of profits attained from playing NI and I_N:

$$\begin{aligned} \pi_1^*(NI) \quad &= \quad \frac{b_2(c_\mathrm{b} - b_1 + d_1)^2}{4[b_2 + d_2(N+1)]^2}; \\ \pi_1^*(I_N) \quad &= \quad \frac{(N+2)^3 b_2 (c_\mathrm{b} - b_{1\mathrm{n}} + d_1)^2}{2[4b_2 + d_2(N+2)^2]^2} - r(b_1 - b_{1\mathrm{n}})^2. \end{aligned}$$

I_N is preferable if $\pi_1(I_N) \geq \pi_1(NI)$. Condition $\pi_1(I_N) \geq \pi_1(NI)$ is satisfied if

$$\frac{(N+2)^3 b_2 (c_\mathrm{b} - b_{1\mathrm{n}} + d_1)^2}{2[4b_2 + d_2(N+2)^2]^2} - \frac{b_2(c_\mathrm{b} - b_1 + d_1)^2}{4[b_2 + d_2(N+1)]^2} \quad \geq \quad r(\Delta b_1)^2$$

or, equivalently, if

$$(r-1)(\Delta b_1)^2 + 2\beta\Delta b_1 - \alpha\beta^2 \quad \leq \quad 0. \tag{7.28}$$

The solution of ineq. (7.28) depends on the value of concavity parameter r. In particular, if $r = 1$, condition (7.28) is satisfied for $\Delta b_1 \in (0, \alpha\beta/2)$. If $r > 1$, it is satisfied for any

$$\Delta b_1 \in \left(0, \frac{\beta}{r-1}\left(\sqrt{1 + \alpha(r-1)} - 1\right)\right].$$

Finally, if $0 < r < 1$, this condition is satisfied for

$$\Delta b_1 \in \left(0, \frac{\beta}{1-r}\left(1-\sqrt{1-\alpha(1-r)}\right)\right] \cup \left[\frac{\beta}{1-r}\left(1+\sqrt{1-\alpha(1-r)}\right), \infty\right).$$

Thus, the equilibrium outcome depends on the R&D cost of innovation and thus on the marginal cost difference Δb_1. The subgame-perfect equilibrium action of firm 1 in stage I is given by I_N for a sufficiently low value of Δb_1 (with the notable exception of the case with $r < 1$ when sufficiently large values of r support this equilibrium, too). By way of contrast, if Δb_1 is too high, then the only action of firm 1 sustainable in a subgame-perfect equilibrium is NI.

Proof of Lemma 7.2.8 Suppose that firm 1 does not innovate in stage I by choosing action NI. The fine and the quota levels chosen by the government in stage II are f_{NI} and $\bar{x}_{NI}$. Given a competitive market in stage III, a representative power generator faces the optimization problem

$$\max_{x_{i\mathrm{b}}, x_{i\mathrm{g}}} [p(x_{i\mathrm{b}} + x_{i\mathrm{g}}) - f_{NI}(\bar{x}_{NI} - x_{i\mathrm{g}}) - C_{\mathrm{b}}(x_{i\mathrm{b}}) - C_{\mathrm{g}}(x_{i\mathrm{g}})], \tag{7.29}$$

Quantities of green electricity produced by each firm and their profits are given by

$$x_{i\mathrm{g}}(NI, f_{NI}) = \frac{c_{\mathrm{b}} - b_1 + f_{NI}}{2b_2}; \tag{7.30}$$

$$\pi_i(NI, f_{NI}, \bar{x}_{NI}) = \frac{(c_{\mathrm{b}} - b_1 + f_{NI})^2}{4b_2} - f_{NI}\bar{x}_{NI}. \tag{7.31}$$

Proof of Lemma 7.2.9 In stage IIIb, firm 1 innovates and announces to offer N licenses in return of a royalty of υ^q per unit of green electricity. The government determines welfare-maximizing quota and fine levels $(\bar{x}_{I_N}, f_{I_N})$, $(\bar{x}_{NI_N}, f_{NI_N})$.

Subgame $\Gamma^q_{I_N,NI}$. Suppose that firms of type 2 reject firm 1's offer. Then firm 1 operates with the new cost function $C_{\mathrm{gn}}(x_{1\mathrm{g}})$ while firms of type 2 continue to operate with the cost function $C_{\mathrm{g}}(x_{2\mathrm{g}})$. Thus, firm 1's profit maximization problem is given by

$$\max_{x_{1\mathrm{b}}, x_{1\mathrm{g}}} [p(x_{1\mathrm{b}} + x_{1\mathrm{g}}) - f_{I_N})(\bar{x}_{NI} - x_{1\mathrm{g}}) - C_{\mathrm{b}}(x_{1\mathrm{b}}) - C_{1\mathrm{gn}}(x_{1\mathrm{g}}) - R(b_1 - b_{1\mathrm{n}})], \tag{7.32}$$

while firm 2's profit maximization problem corresponds to (7.29) with $i = 2$, $f_{NI} = f_{NI_N}$, and $\bar{x}_{NI} = \bar{x}_{NI_N}$. Thus, quantities of green electricity produced by firm 1 and firms of type 2 are respectively given by

$$x_{1\mathrm{g}}(I_N, (f_{NI_N}, f_{I_N}), \upsilon, NI) = \frac{c_{\mathrm{b}} - b_{1\mathrm{n}} + f_{I_N}}{2b_2};$$

$$x_{2\mathrm{g}}(I_N, (f_{NI_N}, f_{I_N}), \upsilon, NI) = \frac{c_{\mathrm{b}} - b_1 + f_{NI_N}}{2b_2},$$

with profits therefore amounting to

$$\pi_1(I_N, (f_{NI_N}, f_{I_N}, \bar{x}_{I_N}), \upsilon^q, NI) = \frac{(c_{\mathrm{b}} - b_{1\mathrm{n}} + f_{I_N})^2}{4b_2} - R(b_1 - b_{1\mathrm{n}}) - f_{I_N}\bar{x}_{I_N};$$

$$\pi_2(I_N, (f_{NI_N}, f_{I_N}, \bar{x}_{NI_N}), \upsilon^q, NI) = \frac{(c_{\mathrm{b}} - b_1 + f_{NI_N})^2}{4b_2} - f_{NI_N}\bar{x}_{NI_N}.$$

Subgame $\Gamma^q_{I_N,I}$. Now suppose that firms of type 2 accept firm 1's offer and have to pay a royalty of υ^q per unit of green electricity produced. Thus, all firms operate with the new cost function $C_{\mathrm{gn}}(x_{\mathrm{g}})$. The profit maximization problems of firm 1 and firms of type 2 are respectively given by

$$\max_{x_{1\mathrm{b}},x_{1\mathrm{g}}} [p(x_{1\mathrm{b}}+x_{1\mathrm{g}})-f_{I_N})(\bar{x}_{I_N}-x_{1\mathrm{g}})-C_{\mathrm{b}}(x_{1\mathrm{b}})-C_{\mathrm{gn}}(x_{1\mathrm{g}})+N\upsilon^q x_{2\mathrm{g}}-R(b_1-b_{1\mathrm{n}})];$$
$$\max_{x_{2\mathrm{b}},x_{2\mathrm{g}}} [p(x_{2\mathrm{b}}+x_{2\mathrm{g}})-f_{I_N})(\bar{x}_{I_N}-x_{2\mathrm{g}})-C_{\mathrm{b}}(x_{2\mathrm{b}})-C_{\mathrm{gn}}(x_{2\mathrm{g}})-\upsilon x_{2\mathrm{g}}],$$

The quantities of green electricity produced are

$$\begin{aligned}
x_{1\mathrm{g}}(I_N,(f_{NI_N},f_{I_N}),\upsilon^q,I) &= \frac{c_{\mathrm{b}}-b_{1\mathrm{n}}+f_{I_N}}{2b_2};\\
x_{2\mathrm{g}}(I_N,(f_{NI_N},f_{I_N}),\upsilon^q,I) &= \frac{c_{\mathrm{b}}-(b_{1\mathrm{n}}+\upsilon^q)+f_{I_N}}{2b_2},
\end{aligned}$$

with profits thus amounting to

$$\begin{aligned}
\pi_1(I_N,(f_{NI_N},f_{I_N},\bar{x}_{I_N}),\upsilon^q,I) &= \frac{(c_{\mathrm{b}}-b_{1\mathrm{n}}+f_{I_N})^2}{4b_2}+N\upsilon^q\frac{c_{\mathrm{b}}-(b_{1\mathrm{n}}+\upsilon^q)+f_{I_N}}{2b_2}\\
&\quad -R(b_1-b_{1\mathrm{n}})-f_{I_N}\bar{x}_{I_N};\\
\pi_2(I_N,(f_{NI_N},f_{I_N},\bar{x}_{I_N}),\upsilon^q,I) &= \frac{(c_{\mathrm{b}}-(b_{1\mathrm{n}}+\upsilon^q)+f_{I_N})^2}{4b_2}-f_{I_N}\bar{x}_{I_N}.
\end{aligned}$$

Firms of type 2 decide in stage IIIb either to reject (NI) or accept (I) firm 1's offer depending on which of their maximum profits attainable in subgames $\Gamma^q_{I_N,NI}$ and $\Gamma^q_{I_N,I}$ is larger. Thus, its subgame-perfect equilibrium actions with respect to the adoption of the new technology are given as follows:

$$\begin{cases} NI & \text{if} \quad \upsilon^q > \upsilon^{q\max} := \sqrt{(c_{\mathrm{b}}-b_1+f_{I_N})^2+4b_2(f_{NI_N}\bar{x}_{NI_N}-f_{I_N}\bar{x}_{I_N})}-(c_{\mathrm{b}}-b_{1\mathrm{n}}+f_{I_N}) \\ I & \text{if} \quad \text{otherwise.} \end{cases}$$

In other words, $\upsilon^{q\max}$ is the highest possible royalty level at which firm of type 2 innovates.

Firm 1's decision in stage IIIb is based on the maximization of its profits with respect to the royalty level υ^q. Notice that firm 1's profit if firms of type 2 accept the offer $\pi_1(I_N,(f_{NI_N},f_{I_N},\bar{x}_{I_N}),\upsilon^q,I)$, is always at least as high as if the offer is rejected as long as $\upsilon^q\in[0,c_{\mathrm{b}}-b_{1\mathrm{n}}+f_{I_N}]$. Moreover, the profit function $\pi_1(I_N,(f_{NI_N},f_{I_N},\bar{x}_{I_N}),\upsilon^q,I)$ attains its maximum w.r.t. υ^q at the royalty level $\upsilon^q=(c_{\mathrm{b}}-b_{1\mathrm{n}}+f_{I_N})/2$. Thus, firm 1's equilibrium offer υ^{q*} in stage IIIb will be given by

$$\upsilon^{*q}=\begin{cases} \upsilon^{q\max} & \text{if } \upsilon^{q\max}<\dfrac{c_{\mathrm{b}}-b_{1\mathrm{n}}+f_{I_N}}{2};\\[2ex] \dfrac{c_{\mathrm{b}}-b_{1\mathrm{n}}+f_{I_N}}{2} & \text{otherwise.}\end{cases}$$

This offer will always be accepted by firms of type 2 in the equilibrium[10]. The quantity of green electricity produced by any firm of type 2 in the subgame starting at node G_2, $x_{2\mathrm{g}}(I_N, (f_{NI_N}, f_{I_N}, \bar{x}_{I_N}, \bar{x}_{NI_N}), I)$, thus amounts to

$$\begin{cases} \dfrac{2(c_b - b_1 + f_{NI_N}) - \sqrt{(c_\mathrm{b} - b_1 + f_{I_N})^2 + 4b_2(f_{NI_N}\bar{x}_{NI_N} - f_{I_N}\bar{x}_{I_N})}}{2b_2} & \text{if } \upsilon^{q\max} < \dfrac{c_\mathrm{b} - b_{1\mathrm{n}} + f_{I_N}}{2}; \\ \dfrac{c_\mathrm{b} - b_{1\mathrm{n}} + f_{I_N}}{4b_2} & \text{otherwise.} \end{cases}$$

Firms' profits in this subgame are therefore given by $\pi_1(I_N, (f_{NI_N}, f_{I_N}, \bar{x}_{I_N}, \bar{x}_{NI_N}), I) =$

$$= \begin{cases} \dfrac{(c_b - b_{1\mathrm{n}} + f_{I_N})^2}{4b_2} - 2Nf_{NI_N}\bar{x}_{NI_N} + (2N-1)f_{I_N}\bar{x}_{I_N} - R(b_1 - b_{1\mathrm{n}}) \\ + \frac{N\left[3(c_\mathrm{b} - b_{1\mathrm{n}} + f_{I_N})\sqrt{(c_\mathrm{b} - b_1 + f_{I_N})^2 + 4b_2(f_{NI_N}\bar{x}_{NI_N} - f_{I_N}\bar{x}_{I_N})} - (c_\mathrm{b} - b_1 + f_{I_N})^2 - 2(c_\mathrm{b} - b_{1\mathrm{n}} + f_{I_N})^2\right]}{b_2} \\ \text{if } \upsilon^{q\max} < \dfrac{c_\mathrm{b} - b_{1\mathrm{n}} + f_{I_N}}{2}; \\ \dfrac{(N+2)(c_\mathrm{b} - b_{1\mathrm{n}} + f_{I_N})^2}{8b_2} - R(b_1 - b_{1\mathrm{n}}) - f_{I_N}\bar{x}_{I_N} \qquad \text{otherwise,} \end{cases}$$

$\pi_2(I_N, (f_{NI_N}, f_{I_N}, \bar{x}_{I_N}, \bar{x}_{NI_N}), I) =$

$$= \begin{cases} \dfrac{\left[2(c_b - b_{1\mathrm{n}} + f_{I_N}) - \sqrt{(c_\mathrm{b} - b_1 + f_{I_N})^2 + 4b_2(f_{NI_N}\bar{x}_{NI_N} - f_{I_N}\bar{x}_{I_N})}\right]^2}{4b_2} - f_{I_N}\bar{x}_{I_N} \\ \text{if } \upsilon^{q\max} < \dfrac{c_\mathrm{b} - b_{1\mathrm{n}} + f_{I_N}}{2}; \\ \dfrac{(c_\mathrm{b} - b_{1\mathrm{n}} + f_{I_N})^2}{8b_2} - f_{I_N}\bar{x}_{I_N} \qquad \text{otherwise.} \end{cases}$$

Proof of Lemma 7.2.10 Now suppose that firm 1 innovates but offers no licenses to its competitors by choosing I_0. In stage II, the government determines the welfare-maximizing quota and fine levels $\bar{x}_{I_0}$, f_{I_0}, $\bar{x}_{NI_0}$, f_{NI_0}. Firm 1's profit maximization problem is thus given by

$$\max_{x_{1\mathrm{b}}, x_{1\mathrm{g}}} [p(x_{1\mathrm{b}} + x_{1\mathrm{g}}) - f_{I_0})(\bar{x}_{I_0} - x_{1\mathrm{g}}) - C_\mathrm{b}(x_{1\mathrm{b}}) - C_\mathrm{gn}(x_{1\mathrm{g}}) - R(b_1 - b_{1\mathrm{n}})], \tag{7.33}$$

while firm 2's profit maximization problem corresponds to (7.29) with $i = 2$, $f_{NI} = f_{NI_0}$, and $\bar{x}_{NI} = \bar{x}_{NI_0}$. Thus, quantities of green electricity produced by firm 1 and firms of type 2 are respectively given by

$$\begin{aligned} x_{1\mathrm{g}}(I_0, f_{NI_0}) &= \frac{c_\mathrm{b} - b_{1\mathrm{n}} + f_{I_0}}{2b_2}; \\ x_{2\mathrm{g}}(I_0, f_{NI_0}) &= \frac{c_\mathrm{b} - b_1 + f_{NI_0}}{2b_2}. \end{aligned}$$

Firms' profits therefore amount to

$$\begin{aligned} \pi_1(I_0, (f_{I_0}, \bar{x}_{I_0})) &= \frac{(c_\mathrm{b} - b_{1\mathrm{n}} + f_{I_0})^2}{4b_2} - f_{I_0}\bar{x}_{I_0} - R(b_1 - b_{1\mathrm{n}}); \\ \pi_2(I_0, (f_{NI_0}, \bar{x}_{NI_0})) &= \frac{(c_\mathrm{b} - b_1 + f_{NI_0})^2}{4b_2} - f_{NI_0}\bar{x}_{NI_0}. \end{aligned}$$

[10] By assumption, firms of type 2 adopt the new technology if indifferent.

Proof of Lemma 7.2.11 Given the decision of firm 1 not to innovate, the government anticipates all firms' optimal quantity decisions in the subgame Γ^q_{NI} and faces the social welfare function

$$\begin{aligned} W_{NI}(f_{NI},\bar{x}_{NI}) &= Q\left(a-\frac{Q}{2}\right)+(N+1)\pi_i(NI,f_{NI},\bar{x}_{NI})-f_{NI}(N+1)x_{\mathrm{g}}(NI,f_{NI})+ \\ &\quad +d_1(N+1)x_{\mathrm{g}}(NI,f_{NI})-d_2[(N+1)x_{\mathrm{g}}(NI,f_{NI})]^2 \\ &= \frac{a^2-c_{\mathrm{b}}^2}{2}+(N+1)\left[\frac{(c_{\mathrm{b}}-b_1+f_{NI})^2}{4b_2}-f_{NI}\bar{x}_{NI}\right] \\ &\quad +(N+1)f_{NI}\left(\bar{x}_{NI}-\frac{c_{\mathrm{b}}-b_1+s_{NI}}{2b_2}\right)+ \\ &\quad +(N+1)d_1\frac{c_{\mathrm{b}}-b_1+f_{NI}}{2b_2}-d_2(N+1)^2\frac{(c_{\mathrm{b}}-b_1+f_{NI})^2}{4b_2^2}. \end{aligned}$$

One can immediately see that both expressions containing the quota levels cancel out. Thus, this welfare function is identical with that in (7.21) with $s_{NI}=f_{NI}$. Consequently, the government maximizes the welfare function with respect to f_{NI} and sets the socially optimal fine level as

$$f^*_{NI} = \frac{b_2d_1-d_2(N+1)(c_{\mathrm{b}}-b_1)}{b_2+d_2(N+1)}, \tag{7.34}$$

while the quota level $\bar{x}_{NI}$ can be deliberately set by the government. The equilibrium quantities and profits are thus given by

$$\begin{aligned} x_{i\mathrm{g}}(NI) &= \frac{c_{\mathrm{b}}-b_1+d_1}{2[b_2+d_2(N+1)]}; \\ \pi^q_i(NI) &= \frac{b_2(c_{\mathrm{b}}-b_1+d_1)^2}{4[b_2+d_2(N+1)]^2}-\frac{b_2d_1-d_2(N+1)(c_{\mathrm{b}}-b_1)}{b_2+d_2(N+1)}\bar{x}_{NI}<\pi_i(NI). \end{aligned}$$

Proof of Lemma 7.2.12 Given the decision of firm 1 to innovate and offer N licenses, the government anticipates the equilibrium outcomes of subgames $\Gamma^q_{I_N,NI}$, $\Gamma^q_{I_N,I}$, as well as that of the royalty bargaining subgame, and faces the social welfare function $W^q_{I_N}(f_{NI_N},f_{I_N},\bar{x}_{NI_N},\bar{x}_{I_N})$ specified below. Since the outcome of the subsequent subgame crucially depends on whether or not condition

$$\upsilon^{q\max} \geq \frac{c_{\mathrm{b}}-b_{1\mathrm{n}}+f_{I_N}}{2} \tag{7.35}$$

is satisfied, the welfare in stage IIb is given as a piecewise defined continuous function. We distinguish two cases, depending on whether condition (7.35) is fulfilled or not.

Case 1: condition (7.35) is satisfied (the 'otherwise' case in stage IIIb). The government maximizes the welfare function

$$\begin{aligned} W^q_{I_N}(f_{NI_N},f_{I_N},\bar{x}_{NI_N},\bar{x}_{I_N}) &= \frac{(a-c_{\mathrm{b}})(a+c_{\mathrm{b}})}{2}+2(N+1)\frac{(c_{\mathrm{b}}-b_{1\mathrm{n}}+f_{I_N})^2}{8b_2} \\ &\quad -R(b_1-b_{1\mathrm{n}})-(N+1)f_{I_N}\bar{x}_{I_N} \\ &\quad +f_{I_N}\left(\bar{x}_{I_N}-\frac{c_{\mathrm{b}}-b_1+f_{I_N}}{2b_2}\right)+Nf_{I_N}\left(\bar{x}_{I_N}-\frac{c_{\mathrm{b}}-b_{1\mathrm{n}}+f_{I_N}}{4b_2}\right) \\ &\quad +d_1(N+2)\frac{c_{\mathrm{b}}-b_{1\mathrm{n}}+f_{I_N}}{4b_2}-d_2(N+2)^2\frac{(c_{\mathrm{b}}-b_{1\mathrm{n}}+f_{I_N})^2}{16b_2^2} \end{aligned}$$

with respect to $f_{NI_N}, f_{I_N}, \bar{x}_{NI_N}, \bar{x}_{I_N}$ and subject to constraint (7.35). Again, since the quota levels can be set exogenously, the welfare function is identical with that in the subsidy case. The socially optimal fine level is given by

$$f^*_{I_N} = \frac{[2b_2N - d_2(N+2)^2](c_\mathrm{b} - b_\mathrm{1n}) + 2(N+2)b_2d_1}{d_2(N+2)^2 + 4b_2}. \tag{7.36}$$

The maximum welfare level attained in this case is therefore

$$W^{*q}(I_N) = \frac{(N+2)^2(c_\mathrm{b} - b_\mathrm{1n} + d_1)^2}{4[(N+2)^2d_2 + 4b_2]} - R(b_1 - b_\mathrm{1n}).$$

Case 2: condition (7.35) is not satisfied. The government maximizes the welfare function

$$\begin{aligned}
&W'^q_{I_N}(f'_{NI_N}, f'_{I_N}, \bar{x}'_{NI_N}, \bar{x}'_{I_N}) = \frac{(c_b - b_\mathrm{1n} + f'_{I_N})^2}{4b_2} - 2Nf'_{NI_N}\bar{x}'_{NI_N} + (2N-1)f'_{I_N}\bar{x}'_{I_N} - R(b_1 - b_\mathrm{1n}) \\
&+\frac{N\left[3(c_\mathrm{b} - b_\mathrm{1n} + f'_{I_N})\sqrt{(c_\mathrm{b} - b_1 + f'_{I_N})^2 + 4b_2(f'_{NI_N}\bar{x}'_{NI_N} - f'_{I_N}\bar{x}'_{I_N})} - (c_\mathrm{b} - b_1 + f'_{I_N})^2 - 2(c_\mathrm{b} - b_\mathrm{1n} + f'_{I_N})^2\right]}{b_2} \\
&+N\left[\frac{\left[2(c_b - b_\mathrm{1n} + f'_{I_N}) - \sqrt{(c_\mathrm{b} - b_1 + f'_{I_N})^2 + 4b_2(f'_{NI_N}\bar{x}'_{NI_N} - f'_{I_N}\bar{x}'_{I_N})}\right]^2}{4b_2} - f'_{I_N}\bar{x}'_{I_N}\right] \\
&+f'_{I_N}\left(\bar{x}'_{I_N} - \frac{c_\mathrm{b} - b_\mathrm{1n} + f'_{I_N}}{2b_2}\right) \\
&+Nf'_{I_N}\left(\bar{x}'_{I_N} - \frac{2(c_b - b_1 + f'_{NI_N}) - \sqrt{(c_\mathrm{b} - b_1 + f'_{I_N})^2 + 4b_2(f'_{NI_N}\bar{x}'_{NI_N} - f'_{I_N}\bar{x}'_{I_N})}}{2b_2}\right) \\
&+d_1\left(\frac{c_\mathrm{b} - b_\mathrm{1n} + f'_{I_N}}{2b_2} + N\frac{2(c_b - b_1 + f'_{NI_N}) - \sqrt{(c_\mathrm{b} - b_1 + f'_{I_N})^2 + 4b_2(f'_{NI_N}\bar{x}'_{NI_N} - f'_{I_N}\bar{x}'_{I_N})}}{2b_2}\right) - \\
&-d_2\left(\frac{c_\mathrm{b} - b_\mathrm{1n} + f'_{I_N}}{2b_2} + N\frac{2(c_b - b_1 + f'_{NI_N}) - \sqrt{(c_\mathrm{b} - b_1 + f'_{I_N})^2 + 4b_2(f'_{NI_N}\bar{x}'_{NI_N} - f'_{I_N}\bar{x}'_{I_N})}}{2b_2}\right)^2
\end{aligned}$$

with respect to $(f'_{NI_N}, f'_{I_N}, \bar{x}'_{NI_N}, \bar{x}'_{I_N})$ and subject to constraint (7.35) reversed with $<$. It can be shown that, as in the subsidy case, $W^{*q}(I_N) > W'^{q*}(I_N)$ for any $N > 0$. Thus, the optimal decision of the government in stage IIb is given by any combination of fines

$$(f^*_{NI_N}, f^*_{I_N}) = \left(f^*_{NI_N}, \frac{[2b_2N - d_2(N+2)^2](c_\mathrm{b} - b_\mathrm{1n}) + 2(N+2)b_2d_1}{d_2(N+2)^2 + 4b_2}\right)$$

where $f^*_{NI_N}$ satisfies inequality

$$\sqrt{(c_\mathrm{b} - b_1 + f^*_{I_N})^2 + 4b_2(f^*_{NI_N}\bar{x}_{NI_N} - f^*_{I_N}\bar{x}_{I_N})} \geq \frac{3}{2}(c_\mathrm{b} - b_\mathrm{1n} + f^*_{I_N}). \tag{7.37}$$

In this case, the government's choice of the quotas $\bar{x}_{NI_N}, \bar{x}_{I_N}$ is constrained by inequality (7.37).

Proof of Lemma 7.2.13 Given the decision of firm 1 to innovate and offer no licenses, the government anticipates the equilibrium outcome of subgame $\Gamma^q_{I_0}$. As in other subgames, welfare maximization is equivalent to the subsidy case. Here we simply state the socially optimal fine level, which is given by

$$f^*_{NI_0} = f^*_{I_0} \quad = \quad \frac{b_2 d_1 - d_2(N+1)(c_{\mathrm{b}} - b_1) - d_2(b_1 - b_{1\mathrm{n}})}{b_2 + d_2(N+1)}, \tag{7.38}$$

which is identical for the innovating and non-innovating firms as in the subsidy case.

Proof of Proposition 7.2.14 Firm 1 anticipates optimal decisions of the government and its competitors in the subsequent stages and thus decides whether or not to innovate (and if so, whether or not to offer licenses), based on the comparison of its maximum attainable profits given the utility-maximizing decisions of other players. In contrast to the subsidy case, the profit functions in the quota case depend on the quotas set by the government. However, as shown above, the quota levels are not determined from welfare maximization but set exogenously[11].

Here, we assume that the quota level $\bar{x}_I$ set for any innovating firm is equal irrespective of its decision about licenses, $\bar{x}_I := \bar{x}_{I_N} = \bar{x}_{I_0}$. Then we can observe that, for any $N > 0$ and any fine level,

$$\pi_1(I_N, (f_{NI_N}, f_{I_N})) > \pi_1(I_0, (f_{NI_0}, f_{I_0})).$$

Thus, firm 1 will never take action I_0 in stage I. The optimal decision of firm 1 depends on the comparison of maximum attainable profits from choosing NI and I_N, respectively:

$$\begin{aligned} \pi_1^*(NI) \quad &= \quad \frac{b_2(c_{\mathrm{b}} - b_1 + d_1)^2}{4[b_2 + d_2(N+1)]^2} - f^*_{NI}\bar{x}_{NI}; \\ \pi_1^*(I_N) \quad &= \quad \frac{(N+2)^3 b_2(c_{\mathrm{b}} - b_{1\mathrm{n}} + d_1)^2}{2[4b_2 + d_2(N+2)^2]^2} - r(b_1 - b_{1\mathrm{n}})^2 - f^*_{I_N}\bar{x}_{I_N}. \end{aligned}$$

I_N is preferable if $\pi_1^*(I_N) \geq \pi_1^*(NI)$. As in the subsidy case, I_N constitutes a subgame-perfect equilibrium strategy for sufficiently low values of Δb_1, namely when the following inequality is satisfied:

$$A(\Delta b_1)^2 + B\Delta b_1 - C \quad \leq \quad 0,$$

or, equivalently, for any

$$\Delta b_1 \in \left(0, \frac{\sqrt{B^2 - 4AC} - B}{2A}\right]. \tag{7.39}$$

By way of contrast, if Δb_1 exceeds the threshold value of $\frac{\sqrt{B^2-4AC}-B}{2A}$, then the only subgame-perfect equilibrium action of firm 1 is NI (Not Innovate). Note, however, that under the quota policy the threshold level of Δb_1 can be influenced by the government as the quota levels are set exogenously.

[11] With the notable exception of stage IIb, in which constraint (7.5) must be satisfied. However, this is the only constraint for the choice of three variables, $f^*_{NI_N}$, $\bar{x}_{NI_N}$, and $\bar{x}_{I_N}$. In other words, for any free choice of both quota levels, there exists a lower bound for $f^*_{NI_N}$

Proof of Proposition 7.3.1 In stage III, competition takes place given firm 1's decision in stage I and the government's decisions in stage II. Notice that subgame $\tilde{\Gamma}_{NI}$ is equivalent to Γ_{NI} with $s_{NI} = \tilde{s}_{NI}$, subgames $\tilde{\Gamma}_{I,NI}$ and $\tilde{\Gamma}_{I,I}$, respectively, to $\Gamma_{I,NI}$ and $\Gamma_{I,I}$ with $(s_{NI_N}, s_{I_N}) = (\tilde{s}_{NI}, \tilde{s}_I)$, and subgame $\tilde{\Gamma}_{I_0}$, to Γ_{I_0} with $(s_{NI_0}, s_{I_0}) = (\tilde{s}_{NI}, \tilde{s}_I)$. The maximum profit levels of firm 1 in these subgames are therefore given by:

$$\begin{aligned}
\pi_1(NI, (\tilde{s}_{NI}, \tilde{s}_I)) &= \frac{(c_{\rm b} - b_1 + \tilde{s}_{NI})^2}{4b_2}; \\
\pi_1(I_0, (\tilde{s}_{NI}, \tilde{s}_I)) &= \frac{(c_{\rm b} - b_{1\rm n} + \tilde{s}_I)^2}{4b_2} - R(b_1 - b_{1\rm n}); \\
\pi_1(I_N, (\tilde{s}_{NI}, \tilde{s}_I)) &=
\end{aligned}$$

$$= \begin{cases} \dfrac{(c_b - b_{1\rm n} + \tilde{s}_I)^2}{4b_2} - R(b_1 - b_{1\rm n}) + \\ +N[(b_1 - b_{1\rm n}) - (\tilde{s}_{NI} - \tilde{s}_I)]\frac{c_b - b_1 + \tilde{s}_{NI}}{2b_2} & \text{if } (b_1 - b_{1\rm n}) - (\tilde{s}_{NI} - \tilde{s}_I) < \dfrac{c_{\rm b} - b_{1\rm n} + \tilde{s}_I}{2}; \\ \dfrac{(N+2)(c_{\rm b} - b_{1\rm n} + \tilde{s}_I)^2}{8b_2} - R(b_1 - b_{1\rm n}) & \text{otherwise.} \end{cases}$$

Under the pre-commitment regime, the government (G) sets the subsidies without any information about the innovation decision of firm 1. Moreover, firm 1 makes its decision whether to innovate or not (and if so, whether to offer licenses) prior to the announcement of the subsidy levels set by the government. Therefore, both decisions can be considered to be made simultaneously and can be modeled as a normal-form game taking place in stages I and II. In this game, firm 1 chooses one of three actions $\{NI, I_N, I_0\}$, while the government determines a pair of subsidies (s_{NI}, s_I).

In a Nash equilibrium of this normal-form game, any equilibrium strategy of a player must belong to the set of best responses to an equilibrium strategy of the other player. The government's best responses (BR_G) to firm 1's actions are equivalent to its actions in stage II in the no pre-commitment case:

$$\begin{aligned}
s^1 &= (s^1_{NI}, s^1_I) := BR_G(NI) = \left\{ \left(\frac{b_2 d_1 - d_2(N+1)(c_{\rm b} - b_1)}{b_2 + d_2(N+1)}, s_I \right) : \ s_I \in \mathbb{R} \right\}; \\
s^2 &= (s^2_{NI}, s^2_I) := BR_G(I_N) = \left\{ \left(s_{NI}, \frac{[2b_2 N - d_2(N+2)^2](c_{\rm b} - b_{1\rm n}) + 2(N+2)b_2 d_1}{d_2(N+2)^2 + 4b_2} \right) : \right. \\
&\qquad \left. s_{NI} \leq \frac{4b_2(b_1 - b_{1\rm n}) - d_2(N+2)^2(c_{\rm b} - b_1) + b_2(N-2)(c_{\rm b} - b_{1\rm n}) + (N+2)b_2 d_1}{4b_2 + d_2(N+2)^2} \right\}; \\
s^3 &= (s^3_{NI}, s^3_I) := BR_G(I_0) = \\
&= \left(\frac{b_2 d_1 - d_2(N+1)(c_{\rm b} - b_1) - d_2(b_1 - b_{1\rm n})}{b_2 + d_2(N+1)}, \frac{b_2 d_1 - d_2(N+1)(c_{\rm b} - b_1) - d_2(b_1 - b_{1\rm n})}{b_2 + d_2(N+1)} \right).
\end{aligned}$$

Firm 1's best responses (BR_1) to s^1, s^2, and s^3 can be derived by observing its profits as functions of subsidy levels given in (7.12), (7.16), and (7.18). Notice that, since the government's best response to I_N is given by $s^2 = (s^2_{NI}, s^2_I)$ as shown above, condition (7.23) cannot be violated in a Nash equilibrium. Therefore, if firm 1 chooses action I_N in stage I it faces the profit function

$$\pi_1(I_N, (\tilde{s}_{NI}, \tilde{s}_I)) = \frac{(N+2)(c_{\rm b} - b_{1\rm n} + \tilde{s}_I)^2}{8b_2} - R(b_1 - b_{1\rm n}).$$

Moreover, since $\pi_1(I_N, (\tilde{s}_{NI}, \tilde{s}_I)) > \pi_1(I_0, (\tilde{s}_{NI}, \tilde{s}_I))$ for any $(\tilde{s}_{NI}, \tilde{s}_I))$, action I_0 is strictly dominated and thus cannot be played in a Nash equilibrium. Hence, action $s^3 = (s^3_{NI}, s^3_I)$ of the government cannot be supported in an equilibrium since it constitutes a best response to the strictly dominated action I_0 only. Action NI is a best response of firm 1 to (s^1_{NI}, s^1_I) if $\pi_1(NI, (s^1_{NI}, s^1_I)) \geq \pi_1(I_N, (s^1_{NI}, s^1_I))$. A rearrangement shows that this condition is satisfied if

$$s^1_I \in \left[-(c_\mathrm{b} - b_{1\mathrm{n}}) \pm \sqrt{\frac{2b_2}{N+2}\left(\frac{b_2(c_\mathrm{b} - b_1 + d_1)^2}{[b_2 + d_2(N+1)]^2} + 4r(\Delta b_1)^2\right)}\right], \tag{7.40}$$

where $\Delta b_1 = b_1 - b_{1\mathrm{n}}$. Therefore, the first set of Nash equilibria is given if player 1 does not innovate and the government chooses (s^1_{NI}, s^1_I) with s^1_{NI} given above and s^1_I satisfying condition (7.40). By an appropriate choice of s^1_I, the government is able to prevent or allow for the occurrence of this equilibrium. Action I_N is a best response of firm 1 to (s^2_{NI}, s^2_I) if $\pi_1(I_N, (s^2_{NI}, s^2_I)) \geq \pi_1(NI, (s^2_{NI}, s^2_I))$. After solving for s^2_{NI}, we obtain the following condition:

$$s^2_{NI} \in \left[-(c_\mathrm{b} - b_1) \pm \sqrt{b_2\left(\frac{b_2(N+2)^3(c_\mathrm{b} - b_{1\mathrm{n}} + d_1)^2}{[4b_2 + d_2(N+2)^2]^2} - 4r(\Delta b_1)^2\right)}\right]. \tag{7.41}$$

The second set of Nash equilibria is given if player 1 innovates and announces to offer N licenses, while the government chooses (s^2_{NI}, s^2_I) with s^2_I given above and s^2_{NI} satisfying condition (7.41).

Proof of Proposition 7.3.2 We have shown in section 7.2.2 that, due to perfect competition, the optimal decisions of the agents in all subgames are equivalent under subsidy and quota-based policies. Therefore, we derive the solution by considering the normal-form game obtained after the truncation of all subgames following the decisions of the government.

In a Nash equilibrium of this normal-form game, any equilibrium strategy of a player must belong to the set of best responses to an equilibrium strategy of the other player. As in the subsidy case, the government's best responses (BR_G) to firm 1's actions are equivalent to its actions in stage II in the no pre-commitment case:

$$\begin{aligned}
f^1 &= (f^1_{NI}, f^1_I) := BR_G(NI) = \left\{\left(\frac{b_2 d_1 - d_2(N+1)(c_\mathrm{b} - b_1)}{b_2 + d_2(N+1)}, f_I\right) : \; f_I \in \mathbb{R}\right\}; \\
f^2 &= (f^2_{NI}, f^2_I) := BR_G(I_N) = \left\{\left(f_{NI}, \frac{[2b_2 N - d_2(N+2)^2](c_\mathrm{b} - b_{1\mathrm{n}}) + 2(N+2)b_2 d_1}{d_2(N+2)^2 + 4b_2}\right) : \right. \\
&\qquad \left. f_{NI} \leq \frac{4b_2(b_1 - b_{1\mathrm{n}}) - d_2(N+2)^2(c_\mathrm{b} - b_1) + b_2(N-2)(c_\mathrm{b} - b_{1\mathrm{n}}) + (N+2)b_2 d_1}{4b_2 + d_2(N+2)^2}\right\}; \\
f^3 &= (f^3_{NI}, f^3_I) := BR_G(I_0) = \\
&= \left(\frac{b_2 d_1 - d_2(N+1)(c_\mathrm{b} - b_1) - d_2(b_1 - b_{1\mathrm{n}})}{b_2 + d_2(N+1)}, \frac{b_2 d_1 - d_2(N+1)(c_\mathrm{b} - b_1) - d_2(b_1 - b_{1\mathrm{n}})}{b_2 + d_2(N+1)}\right).
\end{aligned}$$

Firm 1's best responses (BR_1) to f^1, f^2, and f^3 can be derived by observing its profits as functions of the fine levels. First notice that, since the government's best response to I_N is given by $f^2 = (f^2_{NI}, f^2_I)$ as above, condition (7.23) cannot be violated in a Nash equilibrium. Therefore, if firm 1 chooses action I_N in stage I it faces the profit function

$$\pi_1(I_N, (\tilde{f}_{NI}, \tilde{f}_I)) \quad = \quad \frac{(N+2)(c_{\mathrm{b}} - b_{1\mathrm{n}} + \tilde{f}_I)^2}{8b_2} - R(b_1 - b_{1\mathrm{n}}).$$

Moreover, since $\pi_1(I_N, (\tilde{f}_{NI}, \tilde{f}_I)) > \pi_1(I_0, (\tilde{f}_{NI}, \tilde{f}_I))$ for any $(\tilde{f}_{NI}, \tilde{f}_I))$, action I_0 is strictly dominated and thus cannot be played in a Nash equilibrium. Consequently, action $f^3 = (f^3_{NI}, f^3_I)$ of the government cannot be supported in an equilibrium since it constitutes a best response to the strictly dominated action I_0 only.
Action NI is a best response of firm 1 to (f^1_{NI}, f^1_I) if $\pi_1(NI, (f^1_{NI}, f^1_I)) \geq \pi_1(I_N, (f^1_{NI}, f^1_I))$. A rearrangement shows that this condition is satisfied if

$$f^1_I \in \left[-(c_{\mathrm{b}} - b_{1\mathrm{n}}) + 4b_2\bar{x}_I \pm \frac{\xi\sqrt{2b_2}}{(N+2)[b_2 + (N+1)d_2]}\right]. \tag{7.42}$$

The first set of Nash equilibria is therefore given if firm 1 does not innovate and the government chooses (f^1_{NI}, f^1_I) with f^1_{NI} given above and f^1_I satisfying condition (7.42). By an appropriate choice of f^1_I, $\bar{x}_{NI}$, and $\bar{x}_I$, the government is able to prevent or allow for the occurrence of this equilibrium.
Action I_N is a best response of firm 1 to (f^2_{NI}, f^2_I) if $\pi_1(I_N, (f^2_{NI}, f^2_I)) \geq \pi_1(NI, (f^2_{NI}, f^2_I))$. After solving for f^2_{NI}, we obtain the following condition:

$$f^2_{NI} \in \left[-(c_{\mathrm{b}} - b_{1\mathrm{n}}) + 2b_2\bar{x}_{NI} - \frac{\psi\sqrt{2b_2}}{4b_2 + (N+2)^2 d_2}, -(c_{\mathrm{b}} - b_{1\mathrm{n}}) + 2b_2\bar{x}_{NI} + \frac{\psi\sqrt{2b_2}}{4b_2 + (N+2)^2 d_2}\right]. \tag{7.43}$$

The second set of Nash equilibria is given if firm 1 innovates and announces to offer N licenses, while the government chooses (f^2_{NI}, f^2_I) with f^2_I given above and f^2_{NI} satisfying condition (7.43).

Chapter 8

Conclusion

The conclusion provides a discussion of possible extensions of the research presented in this dissertation.

In Chapters 3 to 5, we elicited citizens' willingness to pay (WTP) for redistribution through a Discrete Choice experiment performed in 2008. The average Swiss citizen is shown to exhibit a negative marginal willingness to pay (WTP) for redistribution. In addition, a very marked status quo bias would have to be overcome by another payment. Several hypotheses concerning possible determinants of this WTP were tested. However, this analysis is subject to several limitations. First, the chosen methodology allows to test hypotheses about the WTP without any confounding supply-side influences. It should be taken seriously, however, that preferences for redistribution might be subject to influence from political institutions as suggested by the Public Choice literature. In particular, voters might use simplified heuristics based on the positions of familiar parties to infer how a policy will affect them and to cast a vote according to their interests. For instance, Schläpfer et al. (2007) in their Contingent Valuation study show that party information significantly affects individuals' attitudes with respect to public spending. Thus, future work should be devoted to a detailed analysis of political preferences in order to identify whether these factors also influence stated WTP for redistribution. This analysis would, however, require addressing the identification problem once again, since the supply of public redistribution is governed by political institutions. Second, the status quo bias found in this paper calls for more detailed analysis. To the extent that it reflects risk aversion, it should induce demand for redistribution - contrary to the results presented here. One possible explanation why it is so high can be the fact that there are some preferences that are not fully formed [see e.g. Stutzer et al. (2007)]. Another possible explanation might be the redistribution illusion, namely the fact that some respondents are not aware of the actual status quo.

Various improvements and extensions can be identified with regard to chapters 6 and 7. In Chapter 6, the analysis of optimal regulation policies in imperfect markets should be extended to consider a Stackelberg oligopoly. One can think of many extensions to Chapter 7. First, the results are obtained in a parameterized model in order to allow for an explicit calculation of the subgame-perfect equilibrium outcomes. An improvement might consist in the introduction of a more general non-parameterized system of cost and demand functions, thereby allowing for a qualitative comparison of the equilibrium outcomes with those from Chapter 6.

A natural extension of 6 and 7 would contain an econometric analysis of data on markets for renewable energy, thereby allowing to test the theoretical predictions derived in both essays. A particularly interesting empirical extension of these two studies would be to elicit consumers' preferences for renewable energy through a stated choice experiment, in particular estimating the optimal energy portfolio from the point of view of households. The methodology to be applied would roughly correspond to the methodology used in Chapters 3 to 5 [see Schneider and Zweifel (2004, 2009) for a stated choice experiment in the field of energy economics].

Bibliography

Abbasi, S. A. and N. Abbasi (2000). The likely adverse environmental impacts of renewable energy sources. *Applied Energy 65*(1–4), 121–144.

Adsera, A. and C. Boix (2002). Trade, Democracy, and the Size of the Public Sector: The Political Underpinnings of Openness. *International Organization 56*(2), 229–262.

Akkoyunlu, S., I. Neustadt, and P. Zweifel (2009). Why Does the Amount of Income Redistribution Differ between United States and Europe? The Janus Face of Switzerland. SOI Working Paper No. 0810, University of Zurich, Socioeconomic Institute.

Alesina, A. and G.-M. Angeletos (2005). Fairness and Redistribution. *The American Economic Review 95*(4), 960–980.

Alesina, A., R. Baqir, and W. Easterly (1999). Public Goods and Ethnic Divisions. *Quarterly Journal of Economics 114*(4), 1243–1284.

Alesina, A. and P. Giuliano (2009). Preferences for Redistribution. Working Paper.

Alesina, A. and E. Glaeser (2004). *Fighting Poverty in the US and Europe: A World of Difference*. Oxford University Press.

Alesina, A., E. Glaeser, and B. Sacerdote (2001). Why Doesn't the United States Have a European-Style Welfare State? *Brooking Papers on Economic Activity 2*, 187–277.

Alesina, A. and E. La Ferrara (2005). Preferences for Redistribution in the Land of Opportunities. *Journal of Public Economics 89*, 897–931.

Alesina, A. and D. Rodrik (1994). Distributive Politics and Economic Growth. *Quarterly Journal of Economics 109*(1), 465–490.

Amundsen, E. and L. Bergman (2004). Green certificates and market power on the Nordic power market. Expert Report for the SESSA Conference "Addressing Market Power and Industry Restructuring for Consumer Benefits", Stockholm, Sweden, 7–8 October 2004.

Amundsen, E. S. and J. B. Mortensen (2001). The Danish Green Certificate System: some simple analytical results. *Energy Economics 23*(5), 489–509.

Amundsen, E. S. and J. B. Mortensen (2002). Erratum to "The Danish Green Certificate System: some simple analytical results". *Energy Economics 24*(5), 523–524.

Andreoni, J. and J. Miller (2002). Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism. *Econometrica 70*(2), 737–752.

Balcells Ventura, L. (2006). Trade Openness and Preferences for Redistribution: A Cross-National Assessment of the Compensation Hypothesis. *Business and Politics 8*(2), 1–50.

Barenboim, I. and L. Karabarbounis (2008). One Dollar, One Vote. Working Paper.

Baumol, W. J. and W. E. Oates (1988). *The Theory of Environmental Policy*. Cambridge: Cambridge University Press.

Beck, J. H. (1994). An Experimental Test of Preferences for the Distribution of Income and Individual Risk Aversion. *Eastern Economic Journal 20*(2), 131–145.

Ben-Akiva, M. E. and S. R. Lerman (1985). *Discrete Choice Analysis*. MIT Press.

Benabou, R. and E. Ok (2001). Social Mobility and the Demand for Redistribution: the POUM Hypothesis. *Quarterly Journal of Economics 116*(2), 447–487.

Benabou, R. and J. Tirole (2006). Belief in a Just World and Redistributive Politics. *Quarterly Journal of Economics 121*(2), 699–746.

Berry, D. (2002). The market for tradable renewable energy credits. *Ecological Economics 42*(3), 369–379.

Berry, T. and M. Jaccard (2001). The renewable portfolio standard: design considerations and an implementation survey. *29*(4), 263–277.

Bhagwati, J. N. (1969). *Trade, Tariffs, and Growth*. Cambridge, Mass.: MIT Press.

Boeri, T., A. Boersch-Supan, and G. Tabellini (2002). Pension Reforms and the Opinions of European Citizens. *The American Economic Review 92*(2), 396–401.

Boeri, T., A. Boersch-Supan, G. Tabellini, K. O. Moene, and B. Lockwood (2001). Would You Like to Shrink the Welfare State? A Survey of European Citizens. *Economic Policy 16*(32), 7–50.

Bower, J., D. W. Bunn, and C. Wattendrup (2001). A model-based analysis of strategic consolidation in the German electricity market. *29*(12), 987–1005.

Brunner, K. and W. H. Meckling (1977). The Perception of Man and the Conception of Government. *Journal of Money, Credit, and Banking 9*(1), 70–85.

Cameron, D. (1978). The Expansion of the Public Economy: A Comparative Analysis. *American Political Science Review 72*, 1243–1261.

Carlsson, F. and P. Martinsson (2003). Design Techniques for Stated Preference Methods in Health Economics. *Health Economics 12*, 281–294.

CEC (1998). Draft Proposal for a Directive of the European Parliament and of the Council on access of electricity from renewable energy sources to the internal market in electricity. 13 October 1998, Commission of the European Communities, Brussels.

CEC (1999a). Proposal for a Directive of the European Parliament and of the Council on the promotion of electricity from renewable energy sources in the internal electricity market. 23 November 1999, Commission of the European Communities, Brussels.

CEC (1999b). Working Paper of the European Commission. Electricity from renewable sources and the internal electricity market. March, Commission of the European Communities, Brussels.

CEC (2001). Directive 2001/77/EC of the European Parliament and of the Council of 27 September 2001 on the promotion of electricity from renewable energy sources in the internal electricity market. Official Journal of the European Commission, L283: 33–40, Commission of the European Communities, Brussels.

CEC (2008a). Proposal for a directive of the European Parliament and of the Council on the promotion of the use of energy from renewable sources. COM(2008) 19 final, 23 Jan 2008, Commission of the European Communities, Brussels.

CEC (2008b). The support of electricity from renewable energy sources. Accompanying document to the Proposal for a Directive of the European Parliament and of the Council on the promotion of the use of energy from renewable sources. Commission Staff Working Document, SEC(2008) 57, 23 Jan 2008, Commission of the European Communities, Brussels.

Checchi, D. and A. Filippin (2004). An Experimental Study of the POUM Hypothesis. *Research on Economic Inequality 11*, 115–136.

Cukierman, A. and A. H. Meltzer (1986). A Theory of Ambiguity, Credibility, and Inflation under Discretion and Asymmetric Information. *Econometrica 54*(5), 1099–1128.

de Tocqueville, A. (1835). *De la Démocratie en Amérique I.* Paris.

Deininger, K. and L. Squire (1996). A New Data Set Measuring Income Inequality. *World Bank Economic Review 10*(3), 565–591.

Denicolò, V. (1999). Pollution–reducing innovations under taxes or permits. *Oxford Economic Papers 51*(1), 184–199.

Dinica, V. (2006). Support systems for the diffusion of renewable energy technologies—an investor perspective. *Energy Policy 34*(4), 461–480.

Dinica, V. and M. J. Arentsen (2003). Green certificate trading in the Netherlands in the prospect of the European electricity market. *Energy Policy 31*(7), 609–620.

Downs, A. (1957). *An Economic Theory of Democracy.* Harper & Row.

Fan, J., W. Sun, and D. Ren (2005). Renewables portfolio standard and regional energy structure optimisation in China. *Energy Policy 33*(3), 279–287.

Fehr, E. and K. Schmidt (1999). A Theory of Fairness, Competition and Cooperation. *Quarterly Journal of Economics 114*, 817–868.

Fehr, E. and K. Schmidt (2006). *Handbook on the Economics of Giving, Altruism and Reciprocity*, Volume 1, Chapter 8, The Economics of Fairness, Reciprocity and Altruism: Experimental Evidence. North-Holland.

Feld, L. P., J. A. Fischer, and G. Kirchgaessner (2007). The Effect of Direct Democratic Institutions on Income Redistribution: Evidence for Switzerland. Working Paper Series in Economics and Finance 689, Stockholm School of Economics.

Fong, C. (2001). Social Preferences, Self-Interest, and the Demand for Redistribution. *Journal of Public Economics 82*, 225–246.

Fong, C., H. Bowles, and H. Gintis (2006). Strong Reciprocity and the Welfare State. *in: Handbook of the Economics of Giving, Altruism and Reciprocity,*. Vol. 2, Ch. 23.

Fong, C. and F. Oberholzer-Gee (2007). Willingness to Pay for Justice: Evidence from an Experiment on Giving to the Poor. Working Paper.

Garret, G. (2000). Globalization and Government Spending around the World. CEACS Estudio Working Paper 2000/155.

Gilens, M. (1999). *Why Americans Hate Welfare: Race, Media, and the Politics of Antipoverty Policy*. University of Chicago Press.

Gruber, J. and D. M. Hungerman (2007). Faith-Based Charity and Crowd Out during the Great Depression. *Journal of Public Economics 91*(5–6), 1043–1069.

Guillaud, E. (2008). Preferences for Redistribution: a European Comparative Analysis. Working Paper.

Hensher, D. A., J. J. Louviere, and J. D. Swait (1999). Combining Sources of Preference Data. *Journal of Econometrics 89*(1–2), 197–221.

Hirschman, A. and M. Rothschild (1973). The Changing Tolerance of Income Inequality in the Course of Economic Development. *Quarterly Journal of Economics 87*(4), 544–566.

Hole, A. R. (2007). A Comparison of Approaches to Estimating Confidence Intervals for Willingness to Pay Measures. *Health Economics 16*, 827–840.

Hungerman, D. M. (2005). Are Church and State Substitutes? Evidence from the 1996 Welfare Reform. *Journal of Public Economics 89*, 2245–2267.

Inman, Robert, P. and D. L. Rubinfeld (1992). Fiscal Federalism in Europe, Lessons from the United States Experience. *European Economic Review 36*(4), 654–660.

Jensen, S. G. and K. Skytte (2002). Interactions between the power and green certificate markets. *Energy Policy 30*(5), 425–435.

Jensen, S. G. and K. Skytte (2003). Simultaneous attainment of energy goals by means of green certificates and emission permits. *Energy Policy 31*(1), 63–71.

Katzenstein, P. J. (1985). *Small States in World Markets: Industrial Policy in Europe*. Cornell University Press.

Kildegaard, A. (2008). Green certificate markets, the risk of over-investment, and the role of long-term contracts. *36*, 3413–3421.

Kinder, D. R. and L. M. Sanders (1996). *Divided by Color: Racial Politics and Democratic Ideals*. University of Chicago Press.

Koster, F. (2008). The Effects of Social and Political Openness on the Welfare State in 18 OECD Countries. *International Journal of Social Welfare 17*, 291–300.

Kreps, D. and J. Scheinkman (1983). Quantity precommitment and Bertrand competition yield Cournot outcome. *Bell Journal of Economics 14*(2), 326–337.

Kreps, D. M. (1990). *A Course in Microeconomic Theory*. Upper Saddle River/NJ: Prentice Hall.

Kuhn, A. (2005). Subjective Evaluations of Wage Inequality and Preferences for Redistribution. Working Paper, University of Zurich.

Lancaster, K. (1971). *Consumer Demand: A New Approach*. Columbia University Press.

Langniss, O. and R. Wiser (2003). The renewables portfolio standard in Texas: an early assessment. *Energy Policy 31*(6), 527–535.

Lauber, V. (2004). REFIT and RPS: options for a harmonised Community framework. *Energy Policy 32*(12), 1405–1414.

Lemming, J. (2003). Financial risks for green electricity investors and producers in a tradable green certificate market. *Energy Policy 31*(1), 21–32.

Lizzeri, A. and N. Persico (2001). The Provision of Public Goods under Alternative Electoral Incentives. *The American Economic Review 91*(1), 225–239.

Lorenzoni, A. (2003). The Italian Green Certificates market between uncertainty and opportunities. *Energy Policy 31*(1), 33–42.

Louviere, J. J., D. A. Hensher, and J. D. Swait (2000). *Stated Choice Methods - Analysis and Application.* Cambridge University Press.

Luce, D. (1959). *Individual Choice Behavior.* Wiley and Sons, New York.

Luttmer, E. F. P. (2001). Group Loyalty and Taste for Redistribution. *Quarterly Journal of Economics 109*(3), 500–528.

Luttmer, E. F. P. and M. Singhal (2008). Culture, Context, and the Taste for Redistribution. Working Paper.

Madlener, R., W. Gao, I. Neustadt, and P. Zweifel (2009). Promoting renewable electricity generation in imperfect markets: price vs. quantity control. FCN Working Paper No. 1/2008, Institute for Future Energy Consumer Needs and Behavior (FCN). July (revised April 2009).

Madlener, R. and I. Neustadt (2010). Renewable energy policy in the presence of innovation: does government pre-commitment matter? SOI Working Paper, Socioeconomic Institute, University of Zurich. forthcoming.

Madlener, R. and S. Stagl (2005). Sustainability-guided promotion of renewable electricity generation. *Ecological Economics 53*(2), 147–167.

Manski, C. F. and S. F. Lerman (1977). The Estimation of Choice Probabilities from Choice Based Samples. *Econometrica 45*(8), 1977–1988.

Matthes, F. C., S. Poetzsch, and K. Grashoff (2005, September). Power generation market concentration in Europe 1996-2004. An empirical analysis. Report No. 2005-012-en, Öko-Institut e.V., Berlin.

McFadden, D. (1974). Conditional Logit Analysis of Quantitative Choice Behavior. *in: Frontiers of Economics,*, 105–142. ed. P. Zarembka, Academic Press, New York.

McFadden, D. (1981). Econometric Models of Probabilistic Choice. *in: Structural Analysis of Discrete Data with Econometric Applications,*, 198–272. ed. by Ch. Manski and D. McFadden, MIT Press.

McFadden, D. (2001). Economic Choices. *The American Economic Review 91*(3), 351–378.

Meltzer, A. H. and S. F. Richard (1981). A Rational Theory of the Size of Government. *Journal of Political Economy 89*(5), 914–927.

Menanteau, P., D. Finon, and M.-L. Lamy (2003). Prices versus quantities: choosing policies for promoting the development of renewable policy. *Energy Policy 31*(8), 799–812.

Merino-Castello, A. (2003). Eliciting Consumers' Preferences Using Stated Preference Discrete-Choice Models: Contingent Ranking versus Choice Experiment. University Pompeu Fabra Economics and Business Working Paper No. 705.

Milanovic, B. (2000). The Median-Voter Hypothesis, Income Inequality, and Income Redistribution: An Empirical Test with the Required Data. *European Journal of Political Economy 16*(3), 367–410.

Milesi-Ferretti, G. M., R. Perotti, and M. Rostagno (2002). Electoral Systems and Public Spending. *Quarterly Journal of Economics 117*, 609–657.

Molnár, G. and Z. Kapitány (2006a). Mobility, Uncertainty and Subjective Well-Being in Hungary. Discussion Paper 2006/5, Institute of Economics, Hungarian Academy of Science.

Molnár, G. and Z. Kapitány (2006b). Uncertainty and the Demand for Redistribution. Discussion Paper 2006/8, Institute of Economics, Hungarian Academy of Science.

Morthorst, P. E. (2001). Interactions of a tradable green certificate market with a tradable permits market. *Energy Policy 29*(5), 345–353.

Morthorst, P. E. (2003). A green certificate market combined with a liberalised power market. *Energy Policy 31*(13), 1393–1402.

Moser, P. (1994). Constitutional Protection of Economic Rights: The Swiss and U.S. Experience in Comparison. *Constitutional Political Economy 5*(1), 61–79.

Muren, A. and S. Nyberg (2005). Young Liberals and Old Conservatives - Inequality, Mobility and Redistribution. CESifo Working Paper No. 1581.

Neustadt, I. (2010). Do Religious Beliefs Explain Preferences for Income Redistribution? Experimental Evidence. SOI Working Paper, forthcoming, University of Zurich, Socioeconomic Institute.

Neustadt, I. and P. Zweifel (2009). Economic Well-Being, Social Mobility, and Preferences for Income Redistribution: Evidence from a Discrete Choice Experiment. SOI Working Paper 0909, University of Zurich, Socioeconomic Institute.

Neustadt, I. and P. Zweifel (2010a). Income Redistribution: How Should the Pie be Divided? SOI Working Paper, forthcoming, University of Zurich, Socioeconomic Institute.

Neustadt, I. and P. Zweifel (2010b). Is the Welfare State Sustainable? Experimental Evidence. SOI Working Paper No. 1003 University of Zurich, Socioeconomic Institute.

Newbery, D. M. (2002). *Privatization, Restructuring, and Regulation of Network Utilities.* Cambridge: MIT Press.

Nielsen, L. and T. Jeppesen (2003). Tradable green certificates in selected European countries—overview and assessment. *Energy Policy 31*(1), 3–14.

Nilsson, M. and T. Sundqvist (2007). Using the market at a cost: How the introduction of green certificates in Sweden led to market inefficiencies. *Utilities Policy 15*(1), 49–59.

Nishio, K. and H. Asano (2006). Supply amount and marginal price of renewable electricity under the renewables portfolio standard in Japan. *Energy Policy 34*(15), 2373–2387.

OECD (1996). Employment Outlook.

Palmer, K. and D. Burtraw (2005). Cost-effectiveness of renewable electricity policies. *Energy Economics 27*(6), 873–894.

Perotti, R. (1996). Growth, Income Distribution and Democracy: What the Data Say. *Journal of Economic Growth 1*(2), 149–188.

Persson, T., G. Roland, and G. Tabellini (1997). Separation of Powers and Political Accountability. *Quarterly Journal of Economics 112*, 310–327.

Persson, T. and G. Tabellini (1994). Is Inequality Harmful for Growth? *The American Economic Review 84*(2), 600–621.

Persson, T. and G. Tabellini (2000). *Political Economics: Explaining Economic Policy*. MIT Press.

Persson, T. and G. Tabellini (2003). *The Economic Effects of Constitutions*. Munich Lectures in Economics. MIT Press.

Piketty, T. (1995). Social Mobility and Redistributive Politics. *Quarterly Journal of Economics 110*(3), 551–584.

Pizer, W. (1999a). Choosing price or quantity controls for greenhouse gases. Climate Issues Brief No. 17, Resources for the Future, Washington D.C.

Pizer, W. A. (1999b). The optimal choice of climate change policy in the presence of uncertainty. *Resource and Energy Economics 21*(3–4), 255–287.

Rader, N. (2000). The hazards of implementing renewable portfolio standards. *Energy & Environment 11*(4), 391–405.

Rainer, H. and T. Siedler (2008). Subjective Income and Employment Expectations and Preferences for Redistribution. *Economics Letters 99*, 449–453.

Ravallion, M. and M. Lokshin (2000). Who Wants to Redistribute? The Tunnel Effect in 1990s Russia. *Journal of Public Economics 76*, 87–104.

Rawls, J. (1999). *A Theory of Justice*. Belknap Press of Harvard University Press.

Razin, A. and E. Sadka (1995). *Population Economics*. MIT Press.

Roberts, K. W. S. (1977). Voting over Income Tax Schedules. *Journal of Public Economics 8*, 329–340.

Rodriguez, F. C. (1999). Does Distributional Skewness Lead to Redistribution? Evidence from the United States. *Economics and Politics 11*(2), 171–199.

Rodrik, D. (1998). Why Do More Open Economies Have Bigger Governments? *Journal of Political Economy 106*(5), 997–1032.

Romer, T. (1975). Individual Welfare, Majory Voting, and the Properties of a Linear Income Tax. *Journal of Public Economics 4*, 163–185.

Ryan, M. (2004). A Comparison on Stated Preference Methods for Estimating Monetary Values. *Health Economics 13*(3), 291–296.

Sáenz de Miera, G., P. del Rio González, and I. Vizcaíno (2008). Analysing the impact of renewable electricity support schemes on power prices: The case of wind electricity in Spain. *36*(9), 3345–3359.

Samuelson, P. A. (1938). A Note on the Pure Theory of Consumer's Behaviour. *Economics 5*(17), 61–71.

Scheve, K. and D. Stasavage (2006a). Religion and Preferences for Social Insurance. *Quarterly Journal of Political Science 1*(3), 255–286.

Scheve, K. and D. Stasavage (2006b). The Political Economy of Religion and Social Insurance in the United States, 1910-1939. *Studies in American Political Development 20 (Fall)*, 132–159.

Schläpfer, F., M. Schmitt, and A. Roschewitz (2007). Competitive Politics, Simplified Heuristics, and Preferences for Public Goods. SOI Working Paper No. 0712 University of Zurich, Socioeconomic Institute.

Schneider, Y. and P. Zweifel (2004). How much internalization of nuclear risk through liability insurance? *Journal of Risk and Uncertainty 29*(3), 219–240.

Schneider, Y. and P. Zweifel (2009). Willingness to pay for avoiding nuclear risks. SOI Working Paper.

Söderholm, P. (2008). The political economy of international green certificate markets. *36*(6), 2051–2062.

Söderholm, P. and T. Sundqvist (2003). Pricing environmental externalities in the power sector: ethical limits and implications for social choice. *Ecological Economics 46*(3), 333–350.

Stutzer, A., L. Goette, and M. Zehnder (2007). Active Decisions and Pro-Social Behavior. Working Paper No. 07-13, Federal Reserve Bank of Boston.

Telser, H. (2002). *Nutzenmessung im Gesundheitswesen.* Kovač, Hamburg.

Tsoutsos, T., N. Frantzeskaki, and V. Gekas (2005). Environmental impacts from the solar energy technologies. *33*(3), 289–296.

Verbruggen, A. (2004). Tradable green certificates in Flanders (Belgium). *Energy Policy 32*(2), 165–176.

Weitzman, M. L. (1974). Prices vs. quantities. *Review of Economic Studies 41*(4), 477–491.

Weitzman, M. L. (1978). Optimal rewards for economic regulation. *The American Economic Review 68*(4), 683–691.

Zweifel, P. (2007). The Theory of Social Health Insurance. *Foundations and Trends in Microeconomics 3*(3), 183–273.

Curriculum Vitae

Ilja Neustadt was born on the 20th of September 1976 in Moscow (Russian Federation). He studied Economics at the Humboldt University of Berlin and London School of Economics and Political Science (LSE). From 2005 to 2010 he worked on his doctoral thesis at the chair of Prof. Dr. Peter Zweifel (Chair for Applied Microeconomics). As a part of his doctoral studies, he successfully completed the Swiss Doctoral Program in Gerzensee in 2007. During his time at the chair of Prof. Dr. Zweifel he supervised various university courses and carried out several projects.